ABSTRACT

Title of Dissertation: SEARCH HISTORY FOR USER SUPPORT IN INFORMATION-SEEKING INTERFACES

Anita Hajnalka Komlódi, Doctor of Philosophy, 2002

Dissertation directed by: Professor Dagobert Soergel
College of Information Studies

Recording search histories, presenting them to the searcher, and building additional interface tools on them offer many opportunities for supporting user tasks in information seeking and use. This study investigated the use of search history information in legal information seeking. Qualitative methods were used to explore how attorneys and law librarians used their memory and external memory aids while searching for information and in transferring to information use. Based on the findings, interface design recommendations were made for information systems.

Results of the study from the limited legal user group presented evidence of the usefulness of search histories and history-based interface tools. Both user manifestations and researcher observations revealed that searchers need historical information in information seeking. Search histories were found to be useful in many user tasks: memory support, search system use, information seeking, information use, task management, task integration, and collaboration. Task integration and collaboration are extensions of traditional information-seeking and use models. Search histories can support users in integrating information across various user tasks and in collaborating with others.

These findings encouraged the design of user interface tools building on search history information: direct search history displays, history-enabled scratchpad facilities, and organized results collection tools were proposed to support users in their information seeking. Interface designs were developed based on the results of the user needs assessment and they were evaluated through participatory design sessions.

The findings are summarized in a search history framework. The framework consists of the following facets: scope of history, context of history, search history data, search history and result management, search history use, and design features. The various facets set dimensions along which search histories and history-based user interface tools can be defined. This framework can guide the study of search histories in other task domains and the design of interfaces for information systems.

SEARCH HISTORY FOR USER SUPPORT IN INFORMATION-SEEKING INTERFACES

by

Anita Hajnalka Komlódi

Dissertation submitted to the Faculty of the Graduate School of
the University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2002

Advisory Committee:

Professor Dagobert Soergel, Chair
Professor Robert B. Allen
Professor Gary J. Marchionini
Professor Catherine Plaisant
Professor Ben Shneiderman
Professor Marilyn D. White

DEDICATION

To my grandmother, Eszter Lantos

ACKNOWLEDGMENTS

First and foremost, I would like to thank my co-advisors, Dagobert Soergel and Gary Marchionini, for their continued support and guidance throughout this project. I am also greatly indebted to the members of my committee for their valuable feedback all through the process: Robert B. Allen, Catherine Plaisant, Ben Shneiderman, and Marilyn D. White.

I am grateful for the generosity of the study participants, attorneys and law librarians in the Washington, DC area, who contributed their time and ideas to the success of this dissertation. I would like to thank T. R. Halvorson for his wonderful collection of interviews with legal reference experts (Halvorson 2000) and for enabling me to use the text of the interviews to complement my data collection. I owe special thanks to West Group for their generous support of the research through many means, and especially for providing access to their legal information service, Westlaw. I would like to thank the research staff of the Joint World Bank-International Monetary Fund Library for their input through interviews in the early phases of this work.

My thanks go to faculty and students at the College of Information Studies and the Human-Computer Interaction Laboratory at the University of Maryland for providing a wonderful community in which to learn and conduct research.

I am also indebted to the individuals who provided support and feedback in various forms: Richard Salter, Doug Oard, Keith Cogdill, Jinsoo Chung, Jinmook Kim, Ben Bederson, and Wayne Lutters. Special thanks to Harry Hochheiser and Gene Chipman for serving as observers in some of the sessions. Input from participants at three doctoral consortia was also much appreciated: ACM SIGCHI 2000, ISIC 2000, and ASIS&T 2000.

This research was partially supported by a Eugene Garfield Doctoral Dissertation Fellowship, a Soros Foundation Graduate Fellowship, and numerous teaching and research assistantships and travel awards of the University of Maryland. Thanks are due to the Department of Information Systems, University of Maryland Baltimore County for their support in the final phases of the work.

Finally, I am greatly indebted to my family and friends for their never-ending love, encouragement, and support without which this dissertation would not have been possible. Special thanks go to Ben Harper, Tamás Komlódi, Józsefné Komlódi, Lajos and Ildikó Iványi, József Komlódi, Angana Baruah, Laura Slaughter, Péter and Enikő Basa, and Mr. and Mrs. Péter Danos.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

> "Record keeping is generally deficient. "All Greeks know what they should do, but only the Spartans do it." You will find a Spartan or two in the group who describe excellent methods of documenting what they do. For them, documentation serves more than merely historical purposes. It is part of research strategy. It enhances analysis and improves results. That is one reason for keeping records. There are two other reasons: for reference on the same project in case it comes back for supplementation or branching, and for reference in future projects. The non-Spartans all acknowledge that better records should be kept of what was searched and how. The records they keep usually are entirely useless to anyone but the person who wrote them. The way the notes are filed does not make them easily retrievable. Researchers periodically ponder whether they waste more time keeping records that might not be needed, or redoing research because they didn't have records in the instances when they might have been useful. Most researchers believe that, on balance, more time is wasted for lack of records. This problem does not arise because searchers fail to appreciate the value of records. It is a result of the burdens and pressures of work settings that do not allow adequate record-keeping."
> (Halvorson 2000, p. 9-10)

1 Introduction

In his introduction to the collection of interviews with expert legal information seekers, T. R. Halvorson above described the importance of keeping records while looking for information. If record keeping is useful, how can we help end user searchers by automatically recording at least some of the search history information? In our continuous search for information, we use many different tools and skills. A quick search on a computer is an easy start, we ask our friends and colleagues, use libraries, consult reference librarians, order books and periodicals, devise complex searches in specialized databases. We interact with people and objects, both physically and virtually while looking for information, and continuously learn about our need, about the sources and information available, and our environment. We integrate new information into our actions and create a great variety in our goals, tools, interactions, and results while seeking information. Throughout the process of searching, especially with complex information-seeking tasks, we must keep track of our progress, strategize, and maintain information for reuse. We use our memory to bridge across different information sources and activities, but human memory is limited and selective. We create external memory aids to support our memory and to help us keep track of our progress, plan steps, and collect information: we take notes, print out and photocopy information, borrow and buy documents. Computer systems can take some of the burden off the searcher by recording searcher actions and other information automatically. This information can be reused to enhance human memory in various ways, it can be presented to the user through action history displays; new user tools (such as a search result collection system) can be developed, and the search system can apply this information to analyze future actions and display similarities or differences. The goal of this dissertation is to understand search histories and their role in human information seeking and to identify potential application areas for history information to enhance information-seeking user interfaces. Building on a theoretical framework developed through interviews, observations and participatory design sessions', it investigates possible improvements to the human-computer interface of information systems to help users take advantage of search history information, and create and use external memory aids building on search history information. Information seeking and searching are used interchangeably in this work, although it is acknowledged that these are often defined differently.

1.1 Research goals and themes

Information seeking skills have become more important in the last few decades as large amounts of easy-to-access information in everyday life became prevalent through electronic means and end users started searching for information in computerized sources. Factors enhancing and supporting information seeking vary from physical tools (print and electronic) to human and electronic intermediaries and specific skills and knowledge. The overall goal of this dissertation is to make searchers' jobs easier in finding information in electronic environments.

The dissertation sets out to examine searchers' behavior in order to identify and describe search history use and areas of potential use. A thorough description of the nature and role of search histories will form a theoretical framework on which to base interface designs. This framework will be developed through several iterations. History information in information-seeking environments can be applied in many different areas. This research aims to identify potential areas of use for automatically and manually recorded history information to enhance information-seeking interfaces.

The following areas of human behavior are likely to benefit from recorded history information and will be examined in the dissertation. Memory support is the most basic area of application for search histories, this information is recorded so that humans do not have to remember it. In order to find and use information, searchers need to be successful in using the search system, the potential of search histories in supporting this area will be explored. The ultimate goal of the process is finding information/searching and using information. These overall tasks have subprocesses that will be studied individually, such as query formulation or relevance decision making. Information seeking is a problem solving task, managing this task, planning and evaluation can also build on recorded search histories. Based on the findings of the user needs assessment and the literature review, user interface tools and design guidelines are outlined to help searchers use history information.

Information seeking as a process is part of the larger task of the user. When searchers look for information using the same computer they use for creating documents or for other tasks, the search system should help seamlessly embed searching into the larger work process context. Recording the history of actions not only in searching, but also in other processes can help create a continuum between the various tasks. The recorded search history can also help customize systems to users' needs by analyzing log of previous actions.

Another dimension of integration is extending or sharing search histories with other users. Recorded histories are good candidates for sharing with others, searchers often record this information in order to share it. Although this topic is not at the center of the dissertation, the implications of sharing search histories are too strong to ignore and are discussed.

The second goal of the dissertation is to provide a foundation for designing improved information-seeking user interfaces that incorporate search history data. Search histories provide a continuity between past, present and future actions through making these more easily available. History information can be utilized in human-computer interfaces in three ways. Explicit search history displays can give users overviews of the whole of the search process, navigation aids between the different step and even tools for further query formulation or information exploration. Search history information can be integrated in other parts of information-seeking interfaces as well. They can enhance displays by showing relationships between steps (e.g.) result lists by showing what items have been returned previously, can contribute to relevance feedback and recommendation systems, and so on. This implicit use of history information needs to be part of any consideration of interface designs building on this information. A third area of application for search histories involves interface tools built on the availability of search histories, or tools provided to complement and further manage search histories. Tools in this category can include features to transfer information from finding to using or tools to help searchers organize results collected.

1.2 Context and problem definition

1.2.1 Information seeking interfaces

Information seeking computer interfaces designed to search large bibliographic databases were first implemented on online database services. To make the searcher's job easier, search interfaces must be improved. In the beginning, technical capabilities and limitations drove the design of these systems; they were intended for professional intermediary users. Later, online public access catalogs appeared with interfaces for end users. The systems of the 1960s and 1970s provided little interactivity between the searcher and the information base; they did not include functionalities to support a dialogue. Search intermediaries were proficient at both using the systems and searching and thus could accommodate the lack of dialog and limited responses from these systems.

Today non-intermediated searching of full-text electronic databases and the Internet has become an everyday activity for many people who are experts in their domain but not in using these systems or searching in general. As programming tools and other interface technologies have advanced, more elaborate functions, including visualizations, can be integrated into search interfaces to provide richer dialogs between the user and the system. The users in this new environment are often left alone with their screen; human intermediaries have disappeared from the process, and computer interfaces must take over the role of the intermediary and provide for effective interaction between the information source and the user. Advanced technologies permit interactivity at the interface; information intermediaries and interface designers need to work together to take advantage of these capabilities to support end-user information seeking.

Information-seeking interfaces support users in many different tasks while searching for information. The interfaces of the 1960s concentrated on query formulation, while in the 1970s database selection, housekeeping functions, history keeping, and help appeared among the options in computerized interfaces. History-keeping functions were among the first functionalities introduced in information-seeking interfaces. Computers can keep track of user and system actions and objects involved in these actions. The fuller the range of information-seeking actions supported by computers, the more meaningful the capture of user and system actions can be.

Information seeking history information is usually presented in the form of a list of query – search results pairs in temporal order. The introduction of full-text and hypertext capabilities, improved information retrieval systems, and higher levels of interactivity of systems all played a role in a much richer variety of information subtasks supported and captured. Simple lists of queries and results will not be satisfactory in keeping track of user tasks and actions. One example of this trend is treating the following of a hyperlink as a form of query, where the result set can be very different from traditional query-based result sets. Viewing and managing these result sets need new tools and their recording create new challenges. More refined models for history information capture and visualization are needed to account for the improved information-seeking systems. One of the most important reasons to save a record of searching is to share it with others.

1.2.2 Amount of information and interactivity in information-seeking interfaces

The amount of information exchanged between the user and the system is one way to characterize the richness of the interaction. Presenting search histories can provide for improved interaction by making this data available in an easily accessible and manageable format. Interaction is based on action and response cycles: The user acts and the computer reacts, and vice versa. Speed, modes of communication, amount of information, organization and depth of knowledge, and its presentation, among other factors, influence the richness of the interactions. The more information the system provides to the user, the richer the interaction can be.

On the other hand, too much information can overload and confuse the user. Presenting overwhelming amounts of history information can take up too much screen real estate from the search system. History information also competes for the searcher’s attention, it should not distract the user from the main task, but help her focus on it. Selecting the right information and presenting it in the right format and amount is crucial in taking advantage of history information without hindering the searcher.

1.2.3 Information seeking scope of dissertation

This dissertation examines the role of search histories in complex information-seeking tasks by domain expert users in the current information-seeking sessions. Information needs and information-seeking tasks vary in their complexity and characteristics. Simple known-item questions and large-scale topical explorations warrant significantly different subtasks and support. History information can be used to support the complex information-seeking tasks where the effort of maintaining search history pays off by enhancing the results. Simple tasks often involve only one or two steps where it is easy to remember context or previous actions, and where users would not invest the additional time and effort to use history displays. In following through multiple sessions it is assumed that these sessions are related by their topical coverage; however, an interesting future direction would be to examine searchers' information-seeking history over longer periods.

The dissertation examines these issues in the context of legal information seeking. The design involves several Westlaw databases. The legal information-seeking context is described in Section 2.3.

1.2.4 Definition of history information and its contextualization for information-seeking tasks

What makes a set of related or unrelated events a history? The dictionary definition of history is the following:

> "a continuous record of important or public events; the study of past events, especially of human affairs; past events, those connected with a person or a thing; an interesting or eventful past."
>
> (Upshall, p. 514)

Although social history is different from histories of human use of computer applications, they are related in some important ways.

(1) In computer application environments, it is possible to automatically record the events of interaction. Anything and everything can be recorded; the definition of what makes an important event to record is highly context-, user-, and task-dependent. One of the goals of this research is to define what are the important events of information seeking worth capturing and reusing.

(2) In the traditional sense of history, the 'study' of past events is the goal. In applying this definition to information-seeking history information, the focus in most cases is not on studying these events but rather using them in new tasks in a meaningful way. In other scenarios, the explicit study of these events will be the goal, e.g. in using histories in teaching or collaborating, in evaluating and planning searches.

(3) The third definition focuses on the organizing principle of the history information; the events will be recorded around a central theme, a searcher, a topic, or an information-seeking session, depending on the context. Creating relationships between the elements of history is an interesting task that cannot be fully accomplished by automatic logging of events alone.

(4) A history is interesting or eventful only to certain people in certain contexts. What is interesting in what context is determined by user characteristics, needs, and tasks. Selecting and defining events, creating a structure in a group of events, assigning value judgments, setting up relationships, and reusing the information all contribute to creating meaningful and helpful histories from events of an information-seeking session.

Information seeking activities involve many different steps from identifying information needs to making relevance judgments and incorporating the new knowledge into old knowledge structures. In physical environments, state is preserved and informs the user about previous actions. In searching for information, the user may have to focus on a different task for a time and come back to searching; in this case notes, documents, location of physical objects, etc. help users restore the state of the search when they come back to continue it. The preservation of state in digital environments is less obvious and more tools are needed to support it.

Do people learn from the past? Does access to our action history change our actions? Do written records change the way society functions from just having oral histories? What does it mean to humans to keep journals? And does providing a search history improve the efficiency of our information seeking? While this dissertation will not answer all of these question, it hopes to create an understanding of how search histories are used and how the information system can better take advantage of search history information to support information seekers.

2 Literature review and conceptual background

A study of the use of search history information in information-seeking interfaces must examine all the elements involved in this context. The literature review, just as the dissertation research as a whole, centers its treatment of the subject around the user. The first three sections examine various processes of the human searcher: Section 2.1 summarizes cognitive processes, more specifically memory processes, as the main role of search history displays is to provide memory aids. Section 2.1 reviews information-seeking processes, with special emphasis on characteristics that influence the use of search histories. Section 2.3 describes the legal information context, with special attention to the issues important to search history use. Finally, section 2.4 turns to systems and an overview of existing search history interfaces to close the chapter. The final section (2.5) summarizes the implications of the literature review findings for search history displays. These findings are highlighted *in italics* all through the text of this chapter.

2.1 Cognitive background. Memory

Cognitive psychology can inform the design of search history interfaces by examining the cognitive processes of users during information seeking. A history information display is an external memory aid that can help users enhance their memory. This section reviews theories of memory and forgetting based on: Solso (1995), Shneiderman (1998), Reed (1992), and Norman (1975). While the review below does not specifically cite these sources in each section, it builds on all four of them to describe cognitive theories of memory and its implications for the use of search histories as memory aids.

2.1.1 Models of memory

In its study of memory, the field of cognitive psychology has developed several models to describe human memory. A short review of these models creates the context for the discussion of memory and history displays. Information processing (IP) theory describes memory in terms of structural elements and defines processes based on these. It describes sensory memory, short-term or working memory, and long-term memory (LTM) stores. Although this dissertation research bases its findings on the IP model of memory, two other models are reviewed, since aspects of their conceptualization of human memory will serve as basis for describing the use of search histories. Search histories support human memory when it fails: the types of forgetting are described in order to predict areas of application for search histories. The level of processing (LOP) theory of memory differs from the information processing theory in that it focuses on processes and then formulates memory systems based on those. It postulates that incoming stimuli are analyzed through deeper and deeper levels of analysis starting from shallow sensory to deep semantic analysis.

Decisions about analysis are made on the basis of the characteristics of the stimuli and the time available. Tulving describes human memory in terms of episodic and semantic memory. Finally, relevant aspects of knowledge representations and learning are discussed. All of these conceptualizations can be helpful in understanding how humans remember events and information while searching, and how improved interfaces can support them. Relevant sections of these theoretical strides are discussed below and guide the data analysis in the results sections.

2.1.2 Information processing model of memory

The information processing theory of memory depicts memory in terms of information flowing between three interrelated stores. External stimuli are transferred from sensory memory to working memory and finally to long-term memory through a series of processes that select and forward information.

Sensory memory preserves the complete sensory impression for a brief period, between 250 milliseconds and four seconds. Only a fraction of the information in the sensory store will be transferred to working memory through a process called attention. We use attention to select a small number of items out of the visual store to transfer to working memory. This selection process is called focalization (Reed, 1992) and is one of the characteristics of attention. *Of simultaneously arriving information, humans are only capable of processing a certain amount.* Selection is based on the importance of each piece of information, how well it fits into human memory, and what is necessary for normal existence of the human. The information selected is transferred to working memory.

Working memory has limited capacity and is characterized by rapid forgetting rates; it stores seven ± two items or chunks of information for about twelve seconds. The information stored in working memory can be acoustic, visual, semantic, and sensory features. It is the only memory store with any conscious processing ability. The sensory store allows some very low level processing. It serves as a transitory store that can hold a limited amount of information for a limited time used in the production of responses. *Any information that is not worked on or transferred into LTM is lost from working memory.* For example, searchers take advantage of their working memory store when comparing the name of the judge in two documents in two consecutive screens viewed in a brief time period. They have to remember a name for the period of time it takes to traverse from one screen to the next. While they may not remember the name of the judge some time after the search ends, they remember it long enough to compare it to the judge in the second document. Items in working memory are lost if they are not being worked on or kept active. Forgetting from working memory is influenced by interference, described in Section 2.1.3.

Chunking, organizing units of information into groups or clusters, can help working memory performance. It is easier to remember information organized by some rule. To support users' working memory skills, systems can arrange information in meaningful chunks to make remembering and using it easier. If information cannot be chunked into meaningful units, it is harder to remember. Chunking is an important issue for history displays. *If users are allowed to group and organize their past actions it may help them understand and remember them better*. Search history systems can automatically chunk some of the information based on rules and analysis of the events. This can be later revised and augmented by users. Naming these chunks is just as important, as the names define what the user thinks a particular grouping of information is about and what it relates to. The names assigned represent users' thinking of issues and the terms can change over time as the users' interpretations change. Information that is judged to be important can be transferred to long-term memory from working memory through encoding.

Long-term memory (LTM) serves as a repository of knowledge. The capacity of LTM is said to be limitless, and its duration endless. *Information transfers from working memory to LTM through a conscious encoding process, rehearsal or learning.* Knowledge structures in LTM are conceptualized through knowledge representation described in section 2.1.6. Transferring knowledge through learning is also discussed in that section. Based on the information-processing model of memory, *repetition and rehearsal are supported by the ability to easily go back to information encountered earlier*. By providing history tools that bring interaction events closer together through easy access to earlier events may help searchers learn and integrate new information into LTM. We store information in our LTM in an organized way and retrieve it when necessary.

LTM and knowledge stored in it have important roles in information processing. *Decisions about selection, analysis, coding, and storage of stimuli are often based on the user's knowledge stored in the LTM and his current goals.* By enabling users to *represent their goals and parts of their knowledge externally* in the search system user interface, in a way that is integrated with the search screen, *can make it easier to compare new information to these and focus better* than if the searcher must recall the information from memory.

2.1.3 Types of forgetting

A study of memory is important for this dissertation in order to identify avenues of support that can be provided through recording search histories. Memory support is most needed when memory fails, the processes of forgetting prevent us from remembering events of the past. Understanding the factors influencing forgetting can help interface design in trying to eliminate these factors and enhance memory. The processes of forgetting are described here in terms of the IP model of memory. The following three types of forgetting are all relevant to search histories.

1. Loss of information from sensory memory: Failure to attend. If information is not selected from the environment for attention and further processing, it will not transfer to working memory. If the user is not focusing on a certain subsection of the environment, that subsection will be lost to further processing. This subsection can be a section of the screen, which can later turn out to be important.

2. Loss of information from working memory: Failure to encode. Some of the information never transfers from working memory into long-term memory because it is not worked on sufficiently. In the terms of LOP theory described below, this information is not processed deeply enough and is more easily forgotten.

3. Loss of information from LTM due to interference and retrieval cue failure. The decay theory described forgetting as the process of losing information not used or rehearsed over time. This theory offers a very intuitive solution, but is criticized for not accounting for interference. In an experiment by Jenkins and Dallenbach (1924), subjects were asked to memorize a series of nonsense syllables. One group of the subjects went to sleep after memorizing, while others went on with their daily activities. Both groups of subjects were asked to recall the syllables at the same time intervals. The subjects who went to sleep had better recall scores, suggesting that not only decay over time, but interference by other actions and information influences recall. Findings like these led to the theory of interference. The theory of interference in forgetting considers the influence of intervening events between the coding of information and its recall. Retrieval cues are association between memory elements that help their recall. The memory can still exist but cannot be recalled, because the appropriate cues are missing. Sufficient cues are coded with the information at the time of encoding, and if these are not remembered, the information will not be recalled.

2.1.4 Level of processing model of memory

The LOP model of memory was created in reaction to the IP model; it emphasizes the importance of processes and defines structures in terms of the processes. It speculates that the durability of memories is determined by the level of processing that has been accomplished on them. It differs from the information processing theory in that *rehearsal by itself will not help the transfer of information from working memory to LTM, but only if the rehearsal represents deeper levels of processing*, more thorough handling and analysis of the information. Experiments have shown that if the rehearsal is only aimed at maintenance, but does not involve deeper levels of processing, short-term recall is not improved by it. Memory is influenced by how the information was originally encoded; semantically encoded information will be easier to recall than perceptually encoded information. Self-reference of information was also shown to have a strong effect on remembering: if the information concerned the person of the searcher, it was easier to recall it later. A different aspect of personal memories, episodic memory, is described in the next section, 2.1.5. *Providing tools to manage and manipulate search histories and search results helps searchers remember and learn new information.*

2.1.5 Episodic and semantic memory

Episodic and semantic memory types were defined by Tulving (1985). Episodic memory comprises particular experiences dated, timed, and placed in space. An example of episodic memory is a certain visit to a restaurant with specific people. Semantic memory is more abstract, it is a human thesaurus of concepts, terms and their relationships, it is abstract knowledge. In the previous example, semantic memory would be knowledge about the name of the restaurant and the fact that the person goes there from time to time, but no recollection of the specific incident. Episodic and semantic memory systems also differ in the rate of forgetting. Episodic memory is constantly taking in new information and thus it is lost more rapidly. The conceptualization of memory in these two types can be used to described various search history data, as explained in the results section 5.2.2.

2.1.6 Developing knowledge representations and learning

Studies in legal information seeking (Sutton, 1994, Marchionini, 1993) have shown that users, especially *domain expert users such as attorneys, interpret the results of the search as they go along in the information-seeking process, and through interpretation they learn about the topic*. *The intense interaction with the new information can insure its transfer from working memory to LTM and its use in the next steps of searching*. When the searchers are experts in searching, but not the domain, interpretation is left to the end user; the searcher does not work on the information found, and thus forgets it more easily. Different history mechanisms may be useful for these two different groups of users. While domain experts make better use of a scratchpad function in helping them fit the new information into their previous model of the topic, search experts may take advantage of histories as memory aids to record information and collecting results for packaging and delivery at the end of the process. The interpretation while searching for information leads to learning and creating knowledge representations in long-term memory.

Solso (1995) defines knowledge from the cognitive science perspective as " the storage and organization of information in memory" (p. 228). Mental models are one way to describe knowledge representation in long-term memory. Johnson-Laird (1983) defines *mental models as a representation in the mind of real or imaginary situations*. They can represent both physical phenomena and abstract concepts and situations. *Search-history based tools can allow searchers to create external representation of their mental models, interact with them, and record them over time.* The mental model theory of knowledge representation is used in Sutton's (1994) work exploring attorneys' information seeking and learning about legal issues described in section 2.3.

The external and internal factors of forgetting help us understand better the cognitive processes of searching for information. Information seeking involves many earlier steps, exposure to many different units of information in a consecutive sequence. Interference influences forgetting this information, unless the information we found becomes part of our mental model of the topic area we are searching on, it easily is lost from short term memory as we go on to the next step of searching.

The constructivist theory of information seeking, use, and learning depicts *humans constantly building their own world, constructing it dynamically from information continuously encountered. Knowledge is defined by the learner in a very personal, individualized process.* Personalized learning tools can be developed based on the history of activities and building knowledge from information found during searching. Individualized sense-making from information found is also central to the section on information seeking as a process. Kuhlthau and Tama (2001) remark that system design is still insufficient in supporting this individualized information-seeking process. They call for systems designs that "go beyond provision for seeking and gathering to support interpretation and use." (p. 27)

Reflective inquiry (Loh et al. 1997) describes stepping back from an activity to view a history of actions, objects, systems states, and emerging understandings in scientific examination. Reflective inquiry has been promoted in science education to help students learn from small-group scientific enquiry projects by looking back on and analyzing their actions and results. Although reflective enquiry is described in the context of scientific investigation by students, it has direct relevance to this dissertation in that it emphasizes a review of previous events to build an understanding of a complex information space. Loh et al. (1997) described three main reflective cognitive abilities in connection with reflective inquiry:

(1) self-monitoring: strategic planning, reflecting on strategies used, and evaluating outcomes of using strategies;

(2) maintaining goal orientation: keeping goals and subgoals of scientific discovery in mind in the context of the investigation at hand;

(3) developing reflective reasoning and argumentation: coordination of questions and observations.

Loh et al. proposed a software tool, the Progress Portfolio, to support these reflective enquiry subtasks by *exposing transient system states in the process and making these available to the students for reflection.* The tool supports all steps of scientific inquiry from formulating research questions to analyzing and communicating the results. The tool encourages students to formulate research questions by writing them down and later revising them if needed. A Data Camera function is included to help students take pictures of system states for later reflection and manipulation. An annotation tool related to the data objects is provided to allow students to create text and graphical inputs to record how data was generated and what it was interpreted to mean. To support analysis and synthesis, students are allowed to organize their data objects and notes according to their own organizational schemes, create comparisons, and clusters. Chronological records are created by the students through explicitly recording system states. Later they review these. The systems also provides two outlets for communicating the process recorded. The history repository serves students who wish to return to earlier processes, while the presentation tools helps share records. This tool represents a different context, but has many implications for the application of search histories. Supporting reflection in complex information contexts is helpful in user learning, which is an important task of end-user searchers, attorneys.

2.1.7 Implications of theories of memory for history displays

In summary, the various theories of human memory can inform the design of search history interfaces in many ways. Following is a list summarizing factors from the cognitive science literature. First, the implication for search history display design supporting users with the three types of forgetting (discussed above in section 2.1.3) are described. These implications are parallel to the types of forgetting described earlier:

1. History as an aid to attention. Providing search history information to searchers lengthens the amount of time a certain piece of information is available to the user. This increased availability of information from searching (queries, results, documents, groups of documents, steps, etc.) gives users a second chance for attention to these, thus supporting transfer to working memory. Search system screens are often busy and the searcher may not pay attention to all elements of the display. If later reconsideration is necessary, the history record can make it easier for the searcher to return to a previous screen and examine the previously unattended area. Another opportunity for history-based tools to support working memory functions is to provide tools to help chunk information. Presenting information from one screen to the next can also help unload some of the burden from users' working memory.

2. History as an aid to encoding. Search history displays present information-seeking actions and objects in a context that supports the development of structure of the actions and the process by visualizing relationships. Allowing searchers to view and manipulate this structure can help them understand the process better and learn about the system and the information. Providing a clipboard to work with the information gathered in the process of information seeking helps searchers interact with the semantics of new information and thus encode this information into LTM (deeper levels of processing).

3. History as an aid to retrieval (extension of LTM). Search history-based user interface tools can help users record information externally, thus removing burden from human memory processes. Information that will not be learnt to keep in LTM, but will be needed longer than the duration of STM can be recorded through various interface tools partially automatically by the system, partially through user input. In addition to providing information needed, reminder cues to other information encoded in LTM can help people retrieve information from memory.

One of Shneiderman's (1998) eight golden rules of interface design calls designers to reduce working memory load for the user through user interface techniques that reduce the complexity of interfaces. Examples of these techniques are the consolidation of multiple screens, or the reduction of window-motion frequency thus reducing the number of times the user needs to remember information from one screen or window to the next. Appropriate training time should also be allowed for the language of the interface and sequences of actions so that the user can commit this information to long-term memory as opposed to having to keep it in working memory. Shneiderman suggests providing online documentation to the command-syntax forms, abbreviations to make this information available to users in an external memory store. Search history information can serve as a variation of this tool: making previous users action sequences available easily can help with interpretation of system actions and with learning about the language of the interface, thus enabling users to reduce the amount of information kept in working memory. History information can provide a user-centered view of this information as its organization is driven by user actions.

2.2 Information-seeking processes and the role of search histories

Information seeking is a dynamic and complex process. Information-seekers collect information through interactions with an information system, negotiating a match between their information need and the content of the information system. While searching, information seekers construct their own picture of the world by integrating knowledge extracted from the information found with their existing knowledge. Information seeking is a complex task, it comprises of many steps and requires task domain, system, and searching expertise. (Marchionini 1995) The basic assumption of this research is that search histories can help with this complex task. This section examines literature related to the process of information seeking, with special attention to areas where search histories have a role. Analyzing information seeking as problem solving helps to understand the process from a cognitive viewpoint and identify roles for activity histories. Next, stopping and successive searching are examined, then note-taking as a form of creating external memory aids. Finally, the knowledge and information used in searching is reviewed.

2.2.1 Phases of information seeking

Allen (1991) reviews the literature on cognitive research in information science. He describes a convergence between system design and user studies by designing systems with the cognitive characteristics of users in mind. The cognitive paradigm in IR looks at information seeking from the point of view of the user's cognitive (and even emotional) processes. (Taylor 1991, Kuhlthau 1993, Dervin 1992, Belkin 1990) Search history displays have the potential of supporting userss' memory processes, which is one of the cognitive processes of humans. Thus, information-seeking behavior is examined from the cognitive viewpoint here.

Information seeking is a cognitive problem-solving task. When looking at stages and phases in information seeking, the problem-solving model is a helpful starting point, as it defines a generic model of steps in problem-solving tasks. Models have been developed for both general problem solving and specific information-seeking activities. Many of these models have described *planning and evaluation stages at the end of each problem solving cycle*. These steps involve an evaluation of the steps taken and the related outcomes, and using these for future tasks or to learn about the system and the information available. Presenting explicit search session histories can support these steps by recording details about the search history and making those accessible through overviews of the process.

Problem solving is widely discussed in the cognitive science literature. Solso defines problem solving as "thinking that is directed toward the solving of a specific problem that involves both the formation of responses and the selection among possible responses" (Solso 1995, p. 440) Hayes (1989) described the steps of problem solving as follows:

(1) Identifying the problem

(2) Representation of the problem

(3) Planning the solution

(4) Execute the plan

(5) Evaluate the plan

(6) Evaluate the solution

The last two steps of this problem-solving model involve some kind of examination of the larger process of solving a problem, looking at a series of steps and evaluating their effect and outcome. Solso (1995) emphasizes the importance of the representation of the information in problem solving: "… stressed the importance of representing the problem in a way that will enhance your ability to find a solution." Visual external representation of problems and steps taken to solve them will aid working memory in reconstructing and evaluating events and their effects.

Soergel (1985) and Kuhlthau (1991) describe the steps of information seeking as follows:

	Soergel (1985)	**Kuhlthau (1991)**
(1)	Recognize and state the need	Initiation
		Selection
		Exploration
(2)	Develop the search strategy	Formulation
(3)	Execute the search strategy	Collection
(4)	Review search results	
(5)	Edit search results	Presentation
(6)	Check helpfulness of results	

Table 1. Steps of information seeking. (Soergel 1985, Kuhlthau 1991)

As shown in Table 1, in Soergel's model two other processes, interaction and monitoring, bridge across all the other steps. *During monitoring and interaction, searchers learn new information from the system outputs (e.g. results, term suggestions) and modify their behavior accordingly.* The searcher's actions are governed in part by his picture of the overall of the search process.

Kuhlthau's (1991) six stages describe affective and cognitive aspects of searchers' behavior; she claims that there is a gap between the user's cognitive and affective processes and the system's assumptions about these. The last two stages correspond to the last two stages of the problem-solving process described above. *The next to last stage is collection when users gather information that is related to their topic*; they define, extend, and support the focus of their search. The actions of users in this stage include selecting information and making detailed notes. Kuhlthau looked at emotional states of information seekers as well and found that in this stage positive feelings of confidence increases as uncertainty decreases. *In the last stage, presentation, users create a personalized synthesis of the information.* Organizing strategies are applied to prepare the information found for further use. *Search histories can support any of these phases, but especially the planning, evaluation, collection, and information use phases. In addition to supporting individual phases, search histories can provide a link between the various stages through recording and providing information from previous steps.*

Models of information-seeking behavior often describe the process as cycles of refinement: Need – plan – act – evaluate – need – plan – act… The evaluation step usually involves judging the relevance of documents retrieved. The result of this evaluation influences the next set of actions of the user. Providing easy access to previous result sets can help planning the actions of the user, recognition of previously seen documents and result sets is easier for users then recalling these sets.

2.2.2 System support for the whole information-seeking process

As seen above, *the process of information seeking includes many different steps, planned and unplanned, until an information need is satisfied or the process is aborted*. Early information-seeking interfaces concentrated solely on the entry of query expressions and the display of result functions, slowly expanding to include query formulation and reformulation, search aids, examination of result sets and documents, and integration of results into ongoing work. In the design of systems that match the searcher's information-seeking process, one must support the whole flow of the process. Search histories and history-based tools can provide a continuity of actions from the first to the last stages of information seeking and help the transfer to information use.

Vigil reviews developments in the software interface of bibliographic and document retrieval systems in his 1986 ARIST chapter. As a future direction for software interfaces he describes the importance of easy and quick availability of search histories and expert searchers' *ability to quickly assess the logic of what has been done so far in order to be able to formulate the next steps*. He states that "the searcher's continued effectiveness depends on his ability to reconstruct what he had done". (p. 75) He points out that the ability to do this is related to practice in searching.

Visualizing the whole of the searching process can help users clarify their information need by laying out the starting state, intermediate steps, and end results. It also supports the planning and evaluation stages of these activities by providing an overview as a basis for evaluation. Information-seeking actions and directions may change during a session as influenced by new information encountered through searching, annotated search histories can help users keep track of their plans and directions.

System support for the whole of the search process is a very important but challenging task. The smooth flow of steps described in the earlier models usually flow in parallel and in cycles in real-life information-seeking tasks. Users stop for various reasons, their goals and tasks change along the way, these factors influence the provision of search history tools. These factors are described in the next section.

2.2.3 System support for stages, changes and stopping

Research has looked at changes of attention and focus in the information-seeking process. Findings show that *there are many attention changes during the process*; often the scope and goal of information seeking change as well. These shifts in goals and plans can distract from the original goal and plans, histories and history-based tools can help users keep track of their goals and changes in plans and actions.

Bates' berrypicking model (1989) describes human behavior is searching for information as a series of evolving queries and browsing to find information, as opposed to the earlier, linear process of query statement, results set, and identification of best match. She describes information seeking as a process that starts with one feature of a topic or one reference and then traversing through a variety of sources, with each piece of new information encountered giving the searcher new ideas on how to proceed and thus changing her movements direction through the system. The query shifts, in part or whole, at every step of this process, gradually evolving in the light of the new information encountered. This process combines analytical searching and browsing in a fashion similar to berrypicking:

> "Furthermore, at each stage, with each different conception of the query, the user may identify useful information and references. In other words, the query is satisfied not by a single final retrieved set, but by a series of selections of individual references and bits of information at each stage of the ever-modifying search. A bit-at-a-time retrieval of this sort is here called *berrypicking*. This term is used by analogy to picking huckleberries or blueberries in the forest. The berries are scattered on the bushes; they do not come in bunches. One must pick them one at a time. One could do berrypicking of information without the search need itself changing (evolving), but in this article the attention is given to searches that combine both of these features."
>
> (Bates 1989, p. 409)

Bates describes six strategies searchers use: footnote chasing, citation searching, journal run, area scanning, subject searches in bibliographies and abstracting and indexing service, author searching. She defines proposed interface design criteria for all of these with flexible jumps between results lists, full text document, lists of references and citations. This model can be used to describe legal searching, where citation links are followed as a large part of the search strategy. Flexible movements between lists and full texts, easy navigation between different sets, and links to represent citations are important for these types of searches. In addition to these tools, search history displays can be helpful when searchers use these strategies in bridging over the variety of steps and actions and allowing the user to collect information in a centralized and continuous collection space.

Hearst (1999) suggests that the "user interface for information access should allow users to reassess their goals and adjust their search strategy accordingly" based ojn these shifts in direction. She describes *another need when users encounter a "trigger" that leads them into a new direction, but they will later want to return to their original goal. The system interface should support users in returning to the previous branch and continue down that path.* Reassessing goals and going off on tangents cause changes in the behavior.

Robins (1997) examined the change of focus and attention to different aspects of the information problem by analyzing conversation transcripts between end user and intermediaries in information-seeking sessions. He starts from the statement that *"although a user's actions may be intentional and logical, they are not necessarily linear or appear logical"*. He found that the majority of shifts in conversation between the end user and the intermediary occur while being online in search sessions. The search history will record some of these shifts, especially the ones the are 1) reflected in actions, or 2) explicitly recorded by the user. *Search histories can help users bridge across these shifts.*

In relation to discourse analysis, Robins restates Grosz and Snider's (1986) assessment that segments of a discourse are not necessarily related to one another linearly as they occur temporally. This statement refers to discourse; however discourse among the actors is a good reflection of the internal, cognitive processes going on during searching. If we accept the above statement, we can infer that the temporal order of intentions and actions may not necessarily follow a logical order, and their reorganization in a search history interface needs to be left to the user, to reflect meaningful order.

However, Hert (1996) found that the goals of the users she studied remained relatively stable during the search session. She reported that users actively kept to their goal in information-seeking sessions. The changes that occurred pointed to increased detail in the current goal rather than a change in the goal. She attributed the findings that differed from previous research to several factors; one of those has implications for the current research, it concerns weak relevance and user behavior associated with it. Harter (1992) defines weak relevance as "the hope that a citation points to an information entity which will be judged to be relevant when the entity is retrieved." (p. 513) *Weak relevance is associated with insufficient information available to the user, thus not enabling them to make a full relevance judgement, but triggering their interest enough to follow th path if it is easily accomplished in the system.* Hert claims that the relative stability of goals in her study resulted partly from the fact the participants were searching on online public access catalog (OPAC) system that provides only bibliographic pointer information to users, and not the sought information itself. She suggests that this way she only captured a part of the whole of the information-seeking process at the time of the interaction. She speculates that in a full-text system a larger degree of changes in goals would emerge as a result of increased exposure to full-text, primary information. This implies that users in full text systems are more prone to changing goals or follow tangents.

Spink (1997) found that users attend to partially relevant documents if they are unfamiliar with a topic or in the beginning of a search and that these help them refine their queries and change their goals in searching.

Spink et al. (1998) describe the elicitation purposes intermediaries pursue during mediated information seeking through the analysis of discourse between end users and intermediaries. They compare their results to Nahl and Tenopir (1996) and Wu (1993). The most frequently used elicitation referred to concepts and terminology used in searching (50 %), and the second most frequent concerned the relevance of the output (16 %). *This points to the importance of search terms, strategies, and* *relevance judgments in information seeking*. Both of these elicitations tended to occur in strings, that is search intermediaries asked users several questions in a row to inquire about search terms and strategies and output relevance. The fact that in these studies the most elicitations were made while being online points to the importance of having this information available simultaneously with the information-seeking actions.

2.2.4 Note-taking, external problem representation

Searchers themselves create external memory aids even if the system does not provide any. Notes and printouts are paper-based examples of these. The following sections describe studies related to searchers' external memory aid processes.

Spink and Goodrum (1996) describe a study on reference librarians' encoding and external storage (EES) processes during searching using a Boolean information retrieval system. This study will be reviewed here in detail, as its findings are highly relevant to search history displays. Their findings complement cognitive information-seeking behavior models with an analysis of users' EES processes in mediated interactive information seeking. They describe encoding in mediated information seeking as the "human process of creating working notes to assist in the understanding and translation of a user's information problem into a search strategy to retrieve relevant items from an IR system" (p. 684) External storage is defined as the "human process of using recorded notes during an interaction with an IR system" (p. 684)

The study examined artifacts of EES processes, notes taken by intermediaries before and during searching. These notes were content-analyzed and the researchers found that search intermediaries frequently create notes in searching, *an average of 20 entities are created per search*. An entity was defined as unique, independent or separate marks. Entities were categorized into *textual/numerical and graphical groups* with many subgroups. *Some of these notes were created before, some during search sessions.* This raises the question of extending supporting users in the whole of their search process and may indicate the extension of search history displays to include the clarification of the information need and query formulation steps before the actual searching starts.

Spink and Goodrum found that there were personal differences in the content and nature of the notes reference librarians took. They define two ways of encoding working notes: note-taking and user notes. Note-taking is similar to and partly overlapping with what is defined as search history in the current work. It is the "*process of synthesizing, copying or summarizing items or system output retrieved by the IR system* … in addition to or rather than printing the items" (p. 693) Some of this information can be automatically captured and recorded as search history. The user can set the scope of this recording, either as an overall customization of the tool, or at the time of the recording. He can also edit, reorganize, and name this record.

User notes contain additional information created by the user, such as potential search terms, queries, comments, and other similar information. Spink and Goodrum describe the reuse of this information as potentially happening during the same search, in consequent multiple search sessions, or in the output of the search. Some parts of the notes created may not be reused at all. *The findings of this study clearly demonstrate the need for note-taking and annotation tools integrated in search histories to help users record their own interpretation of the events and information found.*

2.2.5 Reducing cognitive load through representing knowledge

Allen (1991) reviews literature on the cognitive theories in information science. He points to two areas where search histories can play an important role in support of users' cognitive processing. One of these is the area of cognitive load, and the other is structuring problem-solving activities. This latter area has been described earlier. Garg-Janardan and Salvendy (1986) examined the role of cognitive engineering in information system design and suggested saving contexts in computer applications as a way to reduce cognitive load. This capability would allow users to view the steps leading to an outcome. They also propose the use of reminders to help users reorient themselves to their recent activities. Proposals of these authors suggest that *providing search histories in information retrieval systems can help eliminate some of the cognitive load imposed on users.*

Studies have examined the different types of knowledge humans use in searching for information. *One way to reduce the cognitive load of the users of any computer application is to present the information they need to complete a task through the interface itself.* Thus presenting menus instead of requiring the users to type commands reduces the cognitive load as users have to recognize and select command instead of having to remember them. Studies of the knowledge and information used in information seeking can point interface designer to specific information to record and display from the searching process.

2.2.6 Implications of information-seeking behavior for history displays

The cognitive nature of information-seeking behavior has many implications for search history designs. Supporting various steps and stages in the process, as well as providing continuity across these steps can be an important role for search histories. The steps of evaluation, planning, information collection, and information use are especially well situated to take advantage of search histories.

Note-taking and annotating activities of users call for interface tools to support these. Searchers create external memory aids even if the system does not support it. Some of these notes can be replaced by automatic history recoding. Incorporating user-created notes with other history information is also important. Representing some of the information needed to execute successful searches can help eliminate some of the cognitive load for users.

2.3 Legal information seeking

The context selected for this research is the legal information field. The domain of legal information can be defined by a limited set of content types, user types, and task types. The context of searches and information need situations are more homogenous than in the case of a general collection. This context will be used to design and test search history interfaces. Some of the characteristics of this domain influence the use of search histories and are described here.

Searching for information is at the heart of the legal field; *thorough legal research is one of the professional responsibilities of legal practitioners* and one of the critical skills lawyers employ on behalf of their clients. It is also part of the ethical obligation of lawyers toward their clients as phrased in ethical codes of the profession. Incompetent legal research may lead to disciplinary action against attorneys or liability to the client for legal malpractice. In less serious cases, a reputation for poor legal research can hurt an attorney's practice.

In describing the legal information seeking, Sutton's (1994) attorney mental model development theory is in the center. Other studies of end-user searching behavior are cited as appropriate when reviewing different aspects of the legal information field.

Sutton criticizes earlier studies for defining relevance as pure topicality. He steps back and defines relevance for legal research as a first step in evaluating legal information. He characterizes relevance as a function of the mental models or conceptual map of the law constructed and maintained by attorneys, "a relevant case is one that plays some cognitive role in the structuring of a legal argument."

Sutton's definition of relevance in the legal information field builds on the event space of the case, placing other similar cases in this space. Sutton describes legal practitioners' cognitive maps of law as having three levels:

(1) base-level modeling of the contours of the event space;

(2) context sensitive exploration of the space and the populating of the relevant subsector;

(3) disambiguation of the subsequent model. He also describes these stages in terms of information-seeking activities, sources, and tools used.

The first level is often accomplished through training, learning about the general issues of an area of the law. The second level, "context sensitive exploration", focuses on a particular issue, how the legal principle has been applied to the facts of the reported case. *Often attorneys start from a seed case and follow links from this in an activity that has been termed "gathering citations" (Stoan 1984), "chaining" (Ellis 1989) or "footnote chasing" and "citation gathering" (Bates 1981). It is a pearl growing activity; collecting increasingly relevant documents based on a starting point.* This is deemed especially important in the legal field because of the use of precedence in similar cases, and the importance of creating the conceptual map of cases with relationships among them. Another search technique used by legal researchers is Shepardizing. Shepard's Citations is a citator tool in legal information that

> "lists later sources that cited earlier sources. A later source is a "citing source"; the earlier source (the one you have located and plan to rely on) is the "cited source." ... Shepard's Citations, published by Shepard's/McGraw-Hill, Inc., are the most commonly used legal citators. Shepard's case citators serve four primary functions:
>
> (1) They provide parallel citations to your cited case.
>
> (2) They trace its history.
>
> (3) They help you determine the treatment of the case by leading you to other cases that have cited it.
>
> (4) They provide references to commentary sources that discuss the case."
>
> (Kunz, 1996, pp. 152-153)

The Westlaw systems used in the dissertation research has a service with very similar goals, it is called Keycite. The terms Keycite and Shepardize are used interchangeably in this dissertation both by the researcher and study participants.

The third level describes the process of disambiguation among the results retrieved by the attorney's searches. *In this process, cases in the event space are evaluated based on their juristic status and treated according to their status. If their jurisdiction does not require their use in the attorney's litigation, they are removed or their influence diminished. The cases are Shepardized in order to examine their current status and the results of this are also taken into consideration when deciding the impact of the cases.* This process is complex and involves investigating the current status of the case through many tools. At the end of this process the mental model is finalized for the time being. Sutton remarks that the three processes are going on in parallel in real-life information-seeking situations. The process described in this dissertation focuses on the second and third levels of mental model building, in which the attorney starts out with a picture of the legal area, explores the conceptual space with the help of this model, updates the model and then uses it to judge new results and update it with new information found. The focus is on these two phases because these involve information seeking to a greater extent than the first level. *Each level in the model builds on the mental model developed in the previous level.* The last two levels also require the attorney to conduct more independent information seeking, as opposed to the directed information behavior in level one mental model development.

Sutton makes several suggestions for the improvement of legal information systems based on his attorney mental model theory. Some of these concern the system interface. His three-tiered model of attorney mental models of the law has strong implications for interface and system design. He suggests that the *system "should embrace a dynamic behavioral model of system users including implementation of algorithms to assist in mental model building at all points along the knowledge continuum from base-level modeling through context-sensitive exploration to model disambiguation"* (p. 199) Thus, interfaces need to support all three levels of model building, the first one being a general topical overview of the event space, "getting a feel for it." This can be realized through providing secondary source materials such as legal treatises and encyclopedias. Simply providing these is not enough, organizing the specific area of law by high-level topics, and introducing secondary authority giving an overview and providing links to primary authority presenting the specifics can support this goal. Providing a structured introduction through a secondary source and links to principal primary authority can support context-specific explorations. Presenting linkages between and among citing and cited cases is an important feature for supporting the second phase, context-sensitive exploration. Supporting disambiguation is the most complex and challenging task. Determining juristic status of a case can be supported by the information retrieval (IR) system up to a point but not in all cases. The author remarks that though the system he describes is an IR system and not an expert system, the lines between IR and decision support systems are deliberately blurred in his study.

2.3.1 Content of legal research

The content of legal research can be broken down into primary and secondary law materials. Primary law or primary authority is the law itself. It includes legal precedent, that is case law and code text originating from different bodies of the government. The American legal system builds on legal precedent; thus, *case law is a very important part of primary law. Legal research, especially in litigation, often involves the collection of general information, news, and other materials used as evidence or background information.*

Legal experts on various areas or points of law write secondary authority. Different publications can be of different detail, ranging from general introductions in encyclopedias to highly detailed treatment of narrow topics in legal annotations. These sources explain and interpret primary law, strongly building on it. Primary and secondary materials are often used in conjunction with each other, *secondary materials provide pointers to primary materials, and primary authority references other primary authority in argumentation.*

Legal information is highly textual, users of this information find reading the information very important in making relevance judgments. As mentioned earlier, legal information is also very context-sensitive. Very often, attorneys look for very specific subject matters represented by the facts of a given case. When examining the text, the context of the text segment is very important in interpreting its meaning and usefulness to an attorney's cause. Visualizing this literature poses many challenges starting from users' resistance to visual displays to the highly context-sensitive nature of the text.

Legal topics are very context-sensitive. Searches are often based on the facts of a case; the relevant primary and secondary materials must match this as closely as possible. These complex patterns must be abstracted and compared from both. Legal materials are also highly interlinked, citing other cases and documents is a central part of legal reasoning. This characteristic of legal content lends itself easily to hypertext applications; *legal information systems have taken advantage of hyperlinking in making these documents more accessible.*

2.3.2 Information-seeking techniques and behavior in the legal domain

In accordance with its varied content, legal research involves the use of many different tools and sources. The availability of full-text information systems, hypertext, and other functionalities suggest that computerized systems would be especially helpful in legal research by bridging different sources and publications, and providing a complete set of both primary and secondary resources, general information, news, and potentially evidentiary information. Links between the different resources and full-text search capabilities provide a new environment for legal research. They enable information-seeking solutions that were not possible before, and enrich the field of legal research in many ways. *In this application of search histories, the ability to collect, store, and organize information from multiple sources is useful.*

However, the affordances of print-based publications, such as better context, ease of referenceability, ease of navigating back and forth between documents without losing context, good readability, are in favor of print-based documents. Another factor benefitting print-based sources is the document-centered nature of the legal information field. Legal publishers organize their information output into well-known publications that lawyers are trained on. Thinking in terms of these publications has strong implications for computerized information systems, where maintaining the context of publications is still not solved. *Careful design of these systems must provide a framework for smooth transfers between documents and databases to avoid disorientation.*

Legal research requires high recall in information retrieval. Although attorneys will not use everything they find, they cannot afford to miss anything important in relevant legal precedent or code. This requirement stands in sharp contrast wit the findings that a free-text legal information retrieval system performed well on precision, but poorly on recall measures. (Blair and Maron 1985, Dabney 1986, Berring 1994) Responses came to this criticism from both large legal information providers: Lexis-Nexis and Westlaw. As Sutton points out, the opponents in this debate represent two different frames of reference when addressing the issue of evaluating legal information systems. The critics of the systems base their judgments on the narrow segment of the success of Boolean operators on full-text information systems. The proponents of these two commercial systems examine the question on a much wider basis: considering the improvement these systems brought over manual legal research.

As discussed earlier, Marchionini et al. (1993) found that domain expert *attorneys researching a topic made on-the-fly interpretations of the results*. This resonates Sutton's description of the attorney mental models of law. The two attorneys in the Marchionini et al. study spent a lot of time reading and interpreting the results. As these documents are highly textual, they take longer to skim through. For a document to be relevant often requires that it contain very specific turns of phrases; finding these requires careful reading. New information learnt from results will strongly influence their immediate search strategy. Search intermediaries, on the other hand, did not interpret results, but collected a core set of documents to deliver to end users to make interpretations.

Based on a review of the literature and preliminary data collection, the following list of the typical tasks in legal research has been compiled:

- Topical exploration search on new topic
- Known-item/topic search to update references
- Search for cases decided by a given judge or other specific attribute
- Complex topic-driven search
- Browsing in context (avoid disorientation)
- Gathering materials
- Identification of a core set of items from multiple searches
- Integration of materials from multiple sources
- Integration of materials into documents

Many of these will reappear in Chapter 5, when examples of these tasks will be discussed along with implications for search history design. Many of these tasks can be supported through search history-tools.

Information gathered through legal research is often included in new documents, such as briefs, decisions and law reviews, giving pointers and references to documents or parts of documents. Closer integration between searching tools and other application is an important design requirement in many areas but particularly in legal research.

Kuhlthau and Tama (2001) described attorneys' information search through the results of interviews. They found that attorneys with complex tasks experience uncertainty, but as opposed to novice users, they welcome this as a sign of creative, challenging, and important tasks. Their complex tasks were better served by paper-based resources than computerized systems, as they were exploring an area unknown to them as opposed to searching for a well-defined topic that can be characterized by a keyword. *The authors emphasized the inability of information systems to provide "just in time" and "just for me" personalized services that support the creative construction of a complex strategy, in our context strategy for arguing a legal case.*

Systems that are flexible enough to allow exploration and personalized construction of meaning are needed to support complex, information-intensive work tasks. The attorneys in the study also had great difficulties in managing their internal office files. *The authors called for "multilayered information systems to provide a range of functions including organizing office files, searching the Internet and handling email, and supporting the construction of a case for trial."* (p. 41) This finding emphasizes the importance of providing integration across the various tasks of the suer through the interface. Interaction histories are well positioned to achieve this goal, as they provide a continuity from one task to the next.

Marshall et al. (2001) described a study of law students preparing for a Moot Court competition where they practice case litigation against other student teams. They examined students' information-seeking and reading activities, including annotation techniques, in order to test and design an e-book technology, XLibris. *They found that law students' information seeking is carried out in many physical contexts; they often change place while looking for and processing information. Distributed resources result in distributed documents they work on.* In looking at the law students search techniques, link- or citation-following is a frequently used technique in legal searching, while text searching is less frequent. *They point out that annotation techniques are taught to law students, and often different annotations are used to prepare a document for different purposes. Re-reading and re-annotations are frequent practice in the legal field; often annotations are overwritten or selected annotations are marked for a second time.* Annotations vary in importance and usefulness.

Documents collected for the Moot Court trial are organized according to the tasks and purposes they will be used for. At the same time, as students got closer to writing documents of their own, their organization schemes become closer and closer to their writing objectives. Organization schemes changed through the process of working with the documents, reorganizing them was a way to conduct work. Organizations were activity-based and changed several times during the study. *The authors suggested flexible organization tools that allow reorganization easily*. Students also often created reminders and plans through annotations on document printouts that they later used to guide their further research. The briefs that were the end products of this process contained a discussion of the main legal issues accompanied by relevant quotes from cases. *The students mentioned that they would have liked an easy way to locate quotes that they saw while reading documents.* Writing briefs can also prompt students to look for more materials based on ideas they got while writing. Marshall et al. found that students changed between activities very frequently, and thus suggested a "document laptop" to serve their needs as opposed to a single e-book device.

The redesign suggestions for XLibris focused on the following areas: navigation, retrieval, annotation, and organization. In navigation, *better backtracking tools were needed for navigating among gathered documents and frequent link-following as a search technique*. Keyword-based search facilities were added to *retrieve gathered documents at the passage level*. To support link following, smooth integration with

Web documents was added to the e-book device. Improvements to the annotation tools included the *ability to easily re-annotate previously marked documents and also the ability to annotate previous annotation without having to go back to the original document*, thus supporting thinking. The authors also added a notebook feature where users could collect clippings annotations. More *flexible organization tools* were also added to the original design: *workspace labeling and divider pages* were introduced. Three areas were identified where the new "document laptop" can support users: *1) immediate access to current legal materials through wireless access and highly portable device, 2) ability to re-retrieve previously seen materials, 3) ability to suspend and resume tasks.*

The results of this study (Marshall et al. 2001) are very important for the current dissertation, as many of the research questions looked at and the results found are similar in the two studies. Although the Marshall et al. approached the problem from the point of view of reading and annotations, while the dissertation started from an information-seeking viewpoint, both studies expanded the focus to other tasks *signaling the importance of task integration across computer applications* and physical workspaces. Marshall et al. felt that their most important finding was the move from a dedicated e-book device for reading exclusively, to an integrated "document laptop" that combines other activities with reading. Both studies found similar activities in terms of looking for information, processing information for reuse, and document writing and other information use. This paper was published while data analysis was carried out in the dissertation research. The review of the Marshall et al. study was deliberately put off to avoid biasing the analysis of the data.

2.3.3 Users and user characteristics

Legal information seeking is a complex task and is often carried out by domain or searching experts. *Attorneys and law librarians are both domain and searching experts, attorneys being domain experts with some searching knowledge and librarians vice versa.* One of the important characteristics of attorneys is that they generally have at least a low level of familiarity with the two legal information systems with largest market penetration: Westlaw and Lexis-Nexis. *They are instructed in legal research in law school*, most of them had free access to these systems in law school and have access to them (usually after participating in training sessions) in their practice. This general level of system expertise is an important feature to rely on when designing interfaces for legal information. Law librarians are usually even more familiar with these systems, as they regularly use them. In addition to system knowledge, legal researchers are instructed in specific sources and strategies used in legal information, thus these are frequently used among legal researchers. *Techniques of record keeping and annotating are usually part this legal research education as well.*

As legal research is part of attorneys' work, it is safe to assume that members of this group have some searching experience, either in automated or print sources. In the study by Marchionini et al. (1993) all four subjects formed “reasonable and generally sophisticated” queries. They recognized the various facets of the problems and used those in searching. Sanderson points out that most law students in their setting were experienced LEXIS and Westlaw users – thus “not information illiterates.”

2.3.4 Implications for design

As stated above, information seeking is a very important task for legal practitioners. They are usually trained in it, and practice it with frequency. Legal searchers are usually domain and searching experts, thus system designers can design more complex interfaces for them. In addition to being instructed in legal searching, annotation and record keeping techniques are also usually taught in law schools, thus user interface tools supporting these functions are necessary.

Legal information seeking applies specific tactics to finding and using information. The legal literature is highly interconnected with many links from primary to primary law and from secondary to primary law. This characteristic of legal literature leads to many citation following, pearl growing search techniques. The pearl growing way of gathering information warrants the use of history information or a results gathering tool that enables the searcher to collect information and documents from multiple sources. Another characteristic of legal research is the need for high recall. This need necessitates tools to insure that no document is “lost” in searching, easy ways to retrieving documents seen previously and saving all and any information in a “safe” place for the searcher.

The process of building a conceptual map of a specific area of law is a dynamic learning activity built through a succession of searches and examination of cases and other documents. Sutton’s (1994) model of attorneys’ mental model building describes three levels. Each level builds on the mental model developed at the previous level, tools that help capture an external representation of this model can support information seeking and use. Interpretation is based on continuous evaluation of the information encountered and comparing this information to the mental representation of the user.

On-the-fly interpretation techniques are also typical in legal searching. Reading or skimming documents while searching creates a need for marking and annotating these and creating suer notes about the new information found. The ultimate goal is often the creation of new documents. Tools to help searchers transform information found into new documents are needed.

The studies summarized in this section defined the need for other features of the interface as well. Re-reading and re-annotating documents was an important need identified, and so were flexible organization and reorganization tools. As far as navigation is concerned, backtracking was considered important. Suspending and resuming tasks was another feature described that can be supported by search histories. Re-retrieving previously seen documents was also reported as a requirement and is also a good candidate for search history support. Several authors reported the need for more personalized systems, and system that provide integration across user tasks through one interface.

2.4 Interaction and search history systems

So far the review has focused on the user - cognitive processes, information seeking in general, and later information seeking in the legal context were described. The focus now shifts to the system. Information-seeking history interfaces are described in the following sections.

In physical environments, characteristics of physical objects, landmarks, wear and tear on objects can help us remember. In computer applications, screens provide the environment for our actions, and these screens represent information we need to remember. Showing direct history displays is one way to make up for the lack of physical reminders. Allowing users to create landmarks, recording object histories embedded in the interaction objects is another way to help mimic wear and tear and thus remind searchers of previous actions or previously encountered information.

2.4.1 Framework for system examples

Many information systems collect search history information in some form for various purposes. Due to the lack of space, the current review will be limited to the description of the application areas addressing problems of information seeking, and illustrate them with a few examples from each area. The following areas are described:

2.4.2 Histories in searching

2.4.3 Histories in browsing

2.4.4 Histories in information visualization

2.4.5 History-related information retrieval system functions

Within each of these categories, we can further differentiate systems that serve one individual user, and those focusing on groups of users to support social interaction. As defined in the next chapter, the focus of the current research is systems capturing the search history of one individual user, but the implications of the availability of this information for collaborative applications are very strong and will be discussed. Systems storing and providing information-seeking histories often do so to support collaboration among users. This collaboration can take many forms, users can share usage information to learn from each other, to coordinate collaborative activities, or to guide other users. History information is often recorded in online communities to help strengthen the social aspects of the community and to leverage the common experience and collective wisdom of its members. These systems include collaborative filtering, recommender systems, path visualization systems in hypertext environments, and the like. The success of these systems often depends on a critical mass of people using them to create useful knowledge, and commitment on the part of its users to enter quality information.

2.4.2 Histories in searching

Even the earliest information retrieval systems provided some kind of history mechanism. As described in Chapter 1, these usually involved the display of query – result set pairs. As an example, Back (1976) integrated search review features in his TIRES system, a management information retrieval system, based on the findings of four previous studies and systems. Timbie and Coombs (1969) collected strong positive feedback on the usefulness of search history functions, and Marcus et al. observed that users wanted to have the capability of keeping track of what they have done. Freeman and Atherton (1968) also noted the need for a tool to help users remember search terms, and Katter and McCarn describe a system with history command to trace the structure of search statements. Many early commercial systems had a history feature that allowed users to recall past search commands and reuse those.

Vickery and Vickery (1993) reviewed search interfaces across many systems. They describe many features of online information retrieval systems, among them the collection of data about users. They define three ways a system can collect data about users: user profiles constructed by the user; query specification acquired from the user; and automatic data collection during searches from the queries. They describe the IR-NLI system that among other things, keeps track of records of past sessions by each user.

Hearst (1999) discusses information-seeking behaviors and strategies in her chapter on information retrieval user interfaces and visualizations. She concludes that user observations suggest the need for the user interface of information retrieval and visualization systems to inform users about what steps have been taken in the past; what choices are available; short-term tactics and long-term strategies; and allow the user to annotate choices made and information found along the way. She points out another trend of reusing history information by incorporating personal preference and usage information into the formulation of queries and use of results. She points out that these functions are not well-supported in current systems.

2.4.2.1 Visual history representations

As an example of presenting personal histories, the LifeLines interface (Plaisant et al. 1996) displaying medical histories is reviewed. This interface presents information along a time line, building of the temporal history of any kinds of events. This presentation style is logical for search history information, a time line of search events, as it takes advantage of the temporal sequence of actions. This tools allows the visualization of information organized around different facets, such as in its application to present patient medical records shown in Figure 1.

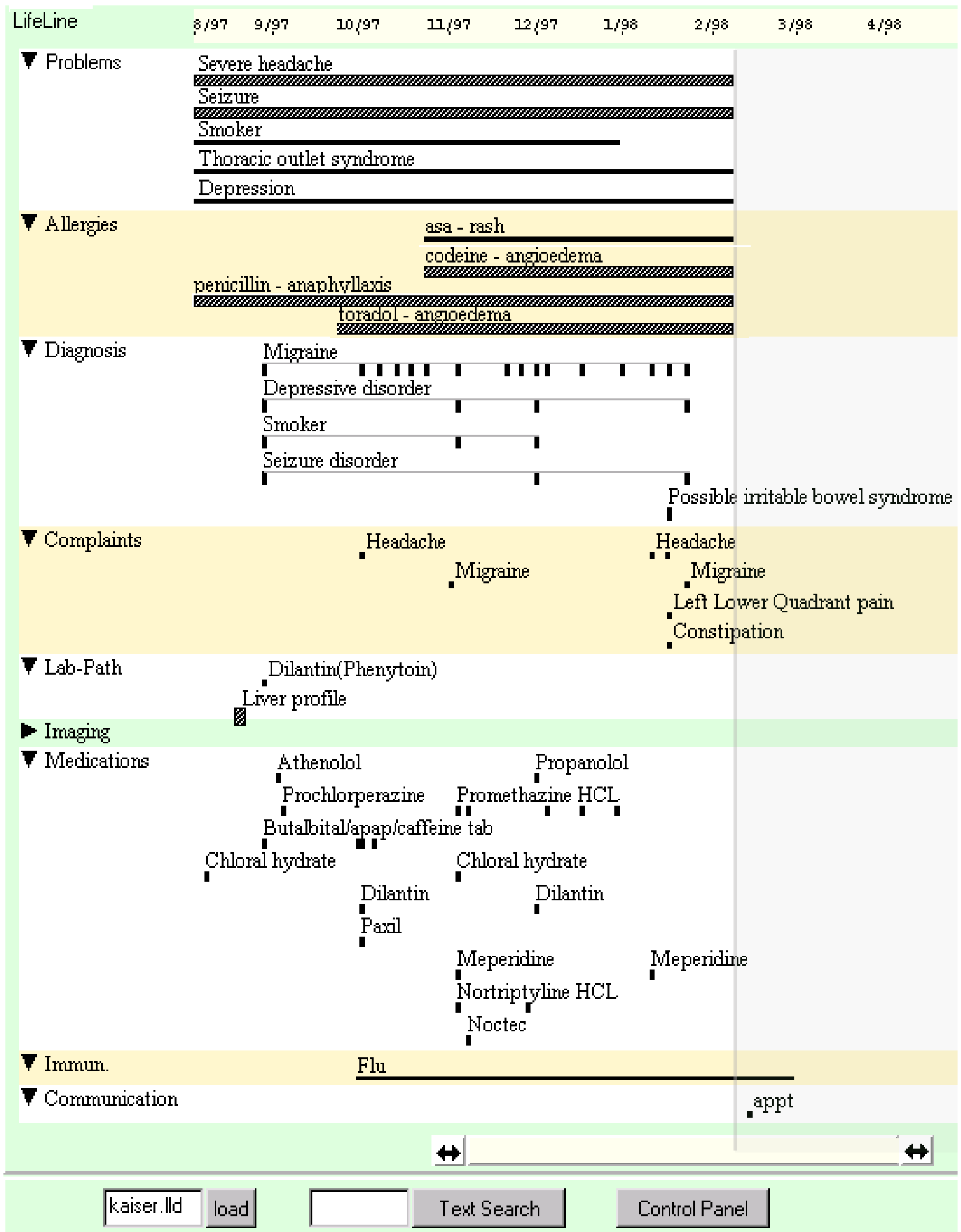


Figure 1. LifeLines interface. Copyright, HCIL UMd, 1998.

The medical record of the patient is represented by a set of lines and events on a zoomable time line. The data in Figure 1. shows events from about two years, along eight facets shown on the left side. These facets can be opened and closed to show an overview of the event or examine details.

Information-seeking processes inherently have a temporal order that is a very useful organizing principle in visual displays. However, temporal order alone does not satisfy all user needs associated with information-seeking session histories. As described earlier, the temporal order may not always be the logical order of events in information seeking. Searchers can jump back and forth between tangents and task avenues. Many of the systems described in this section add other qualities of the events and actions thus enhancing the displays and adding informational content.

Twidale and Nichols (1998) described the Ariadne system developed to support collaboration among users by visualizing search session histories. The system captures command-output (query-result) pairs and displays them to the user as thumbnails of screen shots. The physical layout of the thumbnails reflects both the temporal sequence of actions and the semantic relationships between actions (e.g. choices from menus). The upper part of the screen is divided into three color-coded section. The first section contains thumbnails with query screens, the second results lists, and the third section from the top displays screen shots of individual documents. The capture and display are treated as two separate entities, the visual displays are available separately from searching, in the playback phase. The authors projected the dynamic creation for future versions of Ariadne.

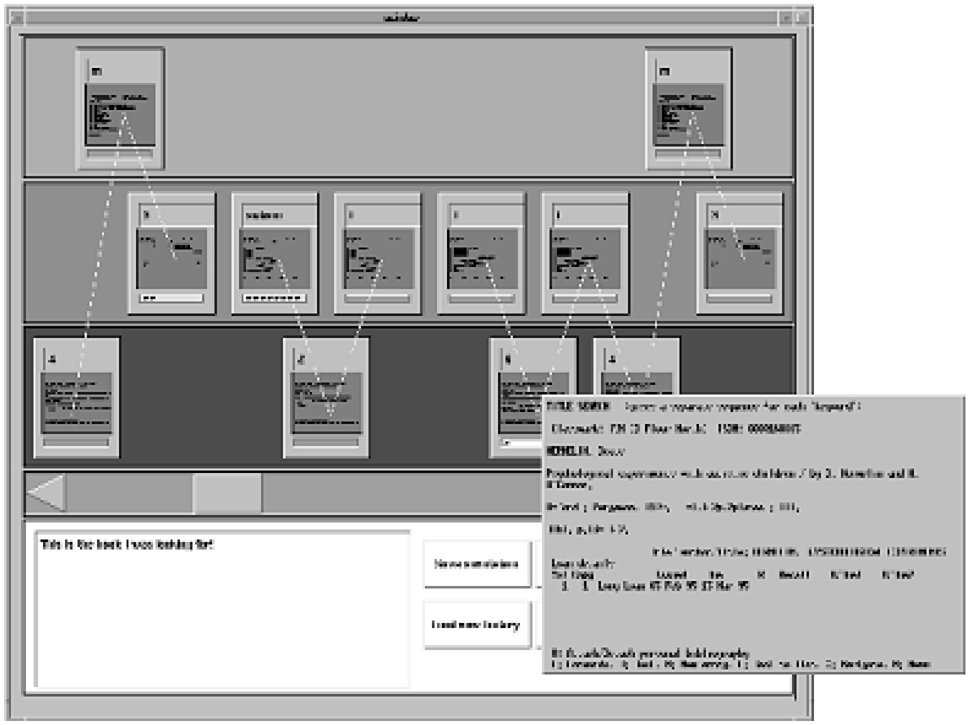


Figure 2. The Ariadne search history interface. Copyright Twidale & Nichols, 1996

Lin et al. (1991) created graphical representations of search patterns by single or multiple searchers in order to study user searching behavior through an analysis of processes and patterns of interactions between users and system. Their goal was to support the recognition of patterns in these interactive environments, to study the process instead of just products of such interactions. Their method offers a good model for defining information-seeking history structures through the following steps:

(1) define a state space of the possible conceptual moves a user can make with a particular system;

(2) capture user keystrokes or mouse clicks during searches;

(3) code these raw data into the state space;

(4) analyze the information-seeking process as a series of conscious moves through the state space.

This last step relates more to the analysis of history data, but was included for completeness. Their method starts out with a formal definition of all possible conceptual moves of a system and analyzes the data based on this. They use this framework to visualize the search process through the conceptual state space of the system – creating a linear sequence of events of a search session. These sessions are visualized on a graph along the possible conceptual moves. Once user logs are recorded and mapped to the conceptual states, animated "movies of the data" are created to compare and analyze user action sequences.

2.4.3 Histories in browsing

2.4.3.1 Hypertext histories

Disorientation is a recognized problem in hypertext environments. Many hypertext studies have shown disorientation effects in these environments. As full-text databases apply hypertext linking to present relationships among the documents, we must consider the effects of disorientation. Hypertext histories have been developed to reduce disorientation. Providing histories in hypertext environments is overlapping, but slightly different from providing search histories. Hypertext structures are relatively stable, the creator of the document sets up links between documents, while in searching many of the links are created on the fly from a large repository of possible links among documents. Items belong to groups in a dynamic manner, memberships change with each search. As HTML technology becomes increasingly dynamic, this will create problems for hypertext histories as well.

In searching, users may start with an information search statements, a specific query. The result of issuing a search command is most often a set of documents or links that vary greatly across the set. Users may browse representations of these documents or the documents themselves, select some for further use or read the documents and finish searching. They may not be satisfied with the set and issue a new search command or modify the previous search. Search histories have many elements on different level: queries, query terms, connectors, limits, result sets, representations of documents, documents, paths or traversal or sequences of actions, relevance judgments, selections, etc. Users may select certain parts of the selected documents and use them in their own work. These elements interact with each other and a record of these interactions can be very useful in informing seeking.

The IBM *Aquí* system is an example where collaborative building of browsing histories creates new structure and takes advantage of the common wisdom of users. It builds on and supports web browsing. The system lets users create links between web pages, records these, and makes them available to other users. In a sense it creates a personalized structure by following user created paths and sharing these with other users. It builds on the experience and recommendation of users to help other users in searching for information. When adding a link to a web page, the user is asked to enter a recommendation on who would like the site linked, an evaluation of how much he liked the site and a self-description of who the user is. Categories of users (e.g. business user, knowledge worker, etc.) are used in both the recommendation and the self-description. Here users make value judgments about sections of their browsing histories and make the ones found most valuable available to other users. As mentioned before, the most useful and beneficial part of histories are most likely the ones that include annotation and interpretations by users.

2.4.3.2 History mechanisms in web browsers

Web browsers are the most often used interfaces for browsing today. Their history mechanisms are influential because of the sheer number of users that encounter and use them, thus forming expectations of history mechanisms in other information-seeking interfaces. Three main tools are used to record history in web browsers: the back button, history lists, and bookmarks/favorites. The back button is a very simple mechanism for backtracking in the sequence of actions building on a stack model of actions. Tauscher and Greenberg (1997) found that users do not have a clear model of how the back button works but still use it very heavily in web navigation. The use of the other two tools is more labor intensive than the back button. To reach the history list, users must open a pull-down menu and then select from the list of items to backtrack. To use bookmarks/favorites, users must record a website for later use. They also have the option of annotating the website and organizing it into folders. In the latest web browsers histories are displayed in a frame next to the main window, so they may be available to users at all times without opening a menu.

Experimental systems have been building alternative ways of capturing and using history information in web browsing. Wexelblat and Maes (1997) describe Footprints, a system for visualizing web browsing histories. They claim that looking at the patterns of many users' activities we can see useful patterns emerge. This web visualization tool shows paths and documents; it does not visualize the document space, but rather the process of users' traversing a site. Thus, the system fits into the groups of systems visualizing a process-oriented view of the document space rather than a more static view.

2.4.4 Histories in visualization environments

2.4.4.1 Visualizing literatures/databases

One should differentiate between visualizing literatures and visualizing the search process. Information visualization often focuses on visualizing the whole of the database and highlighting sections of interest to the user. Visualizing a literature can be described as focusing on a data point-of-view, while visualizing the search process presents the information and the interactions from the users' point of view. Although this differentiation does not always stand, it will be used here to highlight some of the differences between these two. In literature visualization users can manipulate or traverse an overview of an information space (literature). Users can manipulate this space and the system will visualize the reaction of the data to these actions, often without keeping track of users' actions. Visualizing information-seeking session histories could look at the same information from the users' side and visualize the users' actions. This allows for mixed granularity of objects visualized, the inclusion of environment factors in the visualization and higher levels of user control over the display of the information. This way the system would represent the users' cognitive view of the information space and the system's.

Search history visualization can promote better integration of the process from the initial information need and query formulation to the actual use of the document. This may help users in forming a better understanding of the whole of the process, keeping the final objective for which the document will be used when forming the query.

Visualization of data has traditions in the field of numerical data, however legal documents are very textual. The relevance of many legal documents to information needs is often dependent on their relationship to other documents, thus high level overviews of the result space can help users in judging relevance of individual documents and understanding the domain of the problem. Research on the visualization of different literatures, citation networks, and interdisciplinary connections and of the search process has produced many improvements to search interfaces in the research setting. (Card et al. 1999) Visualization techniques must be coupled with good readability and browsability, as legal researchers make relevance judgments based on the exact wording of a statute or fact situation of a case.

Information visualization techniques use graphical tools to present a large amount of information in a dynamic manner. Information visualization is limited here to the visualization of document collections or literatures. Morse (1997) emphasizes the importance for keeping histories in document visualization systems for several reasons, including place keeping and supporting the ability to undo actions. Applying techniques from information retrieval systems to recover and refine intermediate searches in visualization environments would be beneficial. She stresses the importance of retracing steps on a path. Gonzalez (1996) suggests the animation of steps to retrace a path. One of the problems of keeping histories in visualization environments is that of high storage demands for this information. In using visualization environments, the user carries out many sequential actions to arrive at a desired state. To record all information that the user may find useful in the future requires considerable computing power. Morse suggests that in visualization environments it may be useful to keep track of visible landmarks, the granularity of a zoomable display, the state of user-selectable options.

An example of histories in information visualization environments is a history mechanism to keep track of searches for pictures in a zooming environment (Combs and Bederson 1999) This system provides a zooming visualization of the results of a search and saves screen shots as histories of previous searches.

2.4.5 History-related information retrieval system functions

2.4.5.1 Relevance feedback

Information retrieval systems sometimes integrate automatic relevance feedback capabilities to improve retrieval effectiveness. These systems automatically modify the queries based on users' relevance judgments of the information retrieved. A set of documents is retrieved in response to a user query, and these documents are presented to the user for evaluation. The user makes judgments about these documents based on their potential relevance to the query posed. The system uses these judgments in subsequent searches to retrieve documents with the best possible chances of relevance.

Relevance feedback systems record and reuse the history of previous searches and ask the user for additional feedback on these. They realized the augmented histories by asking users to make explicit annotations in the form of relevance judgments to objects previously retrieved. Whenever users are asked to actively record information, they expect high returns for their efforts. Korfhage notes this and two other issues with relevance feedback, that transfer to other information-seeking histories. Some users will want to turn it off; users would get irritated if shown previously judged documents over and over again

2.4.5.2 User profiling and user modeling

User profiles contain information about the user and his information needs, user characteristics and context. The simplest user profiles can consist of a set of query terms, more complex user profiles can describe users' information-seeking practices based on search history information captured through use of the system.

2.4.5.3 Other examples of histories in computerized environments

Hill and Hollan (1994) describe history-enriched digital objects. These objects reflect wear just as physical objects do, enabling users to replay events that make up their use. This information adds to their ease of use in one way or another just as history helps in the case of physical objects. In applications based on this premise, the number or length of events is recorded and displayed to users. These systems implement a very important qualitative aspect of histories: they do not only account for events but also qualify them in more exact terms than physical objects by different characteristics. An example of this feature is Vita Service, an extensive programmable interaction history mechanism that captures time-stamped actions and current context and displays summaries of such as total time spent of various information objects. The history of the digital objects is displayed along with the object, on-demand by the user. The authors also discuss privacy issues related to the capture and sharing of the event information.

Chiueh et al. (1998) describe the Zodiac history-based interactive video authoring system. It provides a flexible video editing environment based on a branching history model of edit operations to organize the authoring process. This tool is a good example of using histories as overview and navigation tools to support strategizing and movement in a system. Zodiac also provides a video object annotation capability, with the interesting feature of having multimedia annotations attached to these objects. The multimedia annotations are attached to moving video objects that are recognized anywhere in the video and the annotation reappear. This content-specific annotation is an interesting tool to have with histories; it can be used to help manage overlaps in information seeking.

2.4.5.4 Use of histories in computer applications generally

Shneiderman (1998) highlights the importance of easy reversal of actions in computer applications. It is important for users to know that their actions are easily reversible. It helps reduce anxiety of using a computer interface and encourage exploration. Shneiderman defines the unit of “undo”-s as a single action, such as a data entry task, or a group of actions, the entry of name and address into an address book. Reversibility of actions is one of the important application areas where history information can be beneficial for users. The ability to go back to any step of the sequence of steps provides a flexible undo mechanisms that makes it easy to reverse whole sequences of actions and try different routes. Well designed search history display can realize the rule of reversibility of actions to a great degree.

Undo tools have been examined in different computer application environments. Leeman (1986) describes undo mechanisms in programming languages and defines models for these. Several authors describe undo mechanisms in collaborative environments. (Ressel et al. 1996; Prakash and Knister 1992, 1994) Berlage describes a selective undo mechanism as a solution to the problem of linear models of events in computer interfaces. Selective undo can copy an isolated command into the current state, thus enabling undoing a single command instead of a whole sequence.

Hanson et al. (1983) examined the use of UNIX commands and found that only a small proportion of all the command available are used; these account for a large percentage of all the usage on the UNIX system. Other studies have shown an important reuse of commands as well, but did not examine the patterns of reuse.

Greenberg and Witten (1988) examined users’ repetition of actions and their use of the history mechanism in the UNIX operating system. They found that the majority of command lines were recurrences of previous actions. They defined a command line as a command, an object, and modifiers. The percentage of recurring commands is the highest for novice users, but the average for all users is still very high, 74%. Another finding of this study concerned the distance of the past action to its recurrence. The previous seven items contributed the majority of recurrences. The most frequent action to be repeated was the second one from the current action.

The authors examined interface techniques to support the reuse of these commands. Directory-sensitive history lists provide context-sensitive histories, showing only command lines referring to files in the directory the user is in. Pruning repetitions from the history lists removes duplicates of command lines to save space, while allowing partial matches would recall previous command that do not exactly match the current one. These techniques can support the reuse of command lines, however they do not translate directly to search histories. In using search histories the goal is often to evaluate and interpret actions and not simple replay them, so all actions need to represented. Partial matches can be useful and context-sensitivity of history displays, as these provide flexibility for users in accessing the history information.

Tauscher and Greenberg (1997) looked at revisitation patterns in web browsing and found that about 60% of all web page visits are revisitations. They found recency to be a reliable predictor of revisitation in the web environment as well. They also confirmed previous research that the Back button is heavily used, but not the history list in web browsers. This has two important implications for the design of search history displays. The first one is a general interface rule: the easier a tool is to use, the more use it will get. The second one considers the expertise and devotion of users to the task. Often web searchers are novices in terms of domain or searching expertise. We often use the web for simple questions, but go to more sophisticated tools for more complex issues. It is hypothesized that users with more complex information problems will be willing to devote more time and effort to their tasks and are more willing to use additional tools for their searching. Tauscher and Greenberg also found that users underused the history command, although they repeated many actions. This finding correlates with the underuse of the history tool found in a study by Yuan (1997) of the information-seeking behavior of law school students in a full-text legal database.

2.4.6 Implications of previous designs

Many of the ideas and solutions reviewed here present useful tools to support information seeking. The focus of the dissertation is search histories.

2.5 Issues identified in the literature review

The literature reviewed here points to the need for recording and displaying search history information in information storage and retrieval systems. The cognitive science background of human memory strengthens the view that forgetting can be supported through recording history information automatically and manually complementing these records. Information seeking behavior of humans and techniques used require many memory processes in each phase separately, especially in planning, evaluation, results collection, and results preparation for reuse. Providing a link across these steps through search histories can help create a continuity and flow of work, and also lead toward integrating searching tasks into larger task environments. Although many current systems provide some history tools, improvements are needed, especially in the search history domain.

3 Search history framework and research questions

This chapter introduces a framework that defines the conceptual space of search histories; this framework was developed iteratively throughout the study. The framework describes user tasks search histories can support, the types of data to record, and interface designs to take advantage of search history information. The chapter defines the focus of the dissertation in this framework. Research questions to guide the dissertation are presented at the end of this chapter based on the definition of the search history framework.

The full framework can be found in Appendix A. The main facets of the framework are:

(1) Scope (Chapter 3)

(2) Context (Chapter 2 & 3)

(3) Search history use (Chapter 5)

(4) Search history and results management (Chapter 5)

(5) Search history data (Chapter 6)

(6) Interface design (Chapter 7)

3.0 Information seeking history framework

The conceptual space of search histories is defined here along several dimensions as shown in Table 2. Some of the dimensions were part of the Wexelblat-Maes interaction history framework, and others were added for the purposes of this dissertation. The following table shows the various dimensions and where they came from:

Active versus passive	Wexelblat-Maes
Rate/form of change	Wexelblat-Maes
Degree of permeation	Wexelblat-Maes
Personal versus social	Wexelblat-Maes
Kind of information	Wexelblat-Maes
Proxemic versus distemic	Wexelblat-Maes
Scope of history:	
Task span	Komlodi
Number of users	Komlodi
Number of systems	Komlodi
Source of data	Komlodi
Automatic vs. manual recording of information	Komlodi
Time span	Komlodi
Context:	
Domain	Komlodi
Organizational context	Komlodi
User characteristics	Komlodi
Complexity of information-seeking task	Komlodi
Access variables	Komlodi

Table 2. Dimension of search history conceptual space.

The goal of the following section is to define the possible range of interaction history systems along several dimensions. Wexelblat and Maes developed an earlier framework for collaborative history systems; their dimensions will be described and complemented with additional criteria for information-seeking contexts.

3.1.1 Interaction history system dimensions: Wexelblat-Maes framework

The Wexelblat-Maes framework will be described as it relates to search history defined in this dissertation. The additional criteria presented reflect the focus of the current research: that of individual history information and the context of information seeking. The criteria presented are orthogonal; they define the design space for these systems. They attempt to describe the options, but the creators of the framework do not claim it to be exhaustive.

3.1.1.1 Active versus passive

This dimension describes the users' relationship to recording history information. History can be recorded passively as users progress through their tasks, without his active participation. On the other hand, the user may stop to think about and actively record history. Actively recorded history is often more useful for users but needs more time and effort to record. The goal is to passively record enough good data to help users while making it easy to add annotations.

The current research aims at designing interfaces that help users use search history information for different tasks. For some tasks, passively recorded information will be sufficient, but in other cases, users may feel it necessary to create annotations and markings. In simple navigation and backtracking tasks, automatically recorded action sequences will enable users to achieve their tasks, but when searchers use histories to help them interpret and integrate information-seeking results user annotation can be of high value.

3.1.1.2 Rate/form of change

As the user progresses through tasks and activities, history information builds up. This accretion of history adds to the richness of the information, but also poses problems. History is "forgotten" after a period, records disappear, people forget, new history information covers up previous information. It also poses a challenge in another way. As the amount of history information grows, tools to manipulate and present it must become increasingly sophisticated to handle the amount of information. Overviews and summaries must be created to help users understand history information. Flexible searching and browsing tools may help in navigating through history.

Recording search histories of a single user through multiple sessions will result in the accumulation of a large amount of information. Techniques will be needed in the interface to manage this information and allow the user to focus in on a specific area of interest. Flexible retrieval and browsing tools must be incorporated into search history displays to achieve this.

3.1.1.3 Degree of permeation

This criterion describes the relationship of the history information to objects of interaction. History can be part of the objects or be kept separately. In the physical world this is often determined by the nature of the object, while in the digital world the history can be recorded and used separately from any object; or be built into any object, depending on the application context. An example of this can be the application of a counter on a web page. If a page has no counter built in it is very hard to tell how many people visit it. We can still find it out from server logs, but this history is stored separately, and is harder to access. If there is a counter on the page, the history becomes part of the object; whenever users visit the page they can learn about its history right away.

Search history information can be displayed through the objects or in separate history displays. An example of integrating history into the interaction object is when a document retrieved in response to a query would show if it was returned previously in another query. Separate history displays can act as navigational time lines, or as more complex displays presenting the conceptual structure of the search process.

3.1.1.4 Personal versus social

History can be recorded for an individual, thus creating a personal history, or it can be created for groups of people helping to create communities. This second group of histories will support collaboration in many fields. Personal histories act on a different level, they augment memory and support problem solving on the level of the individual. As described earlier, the current project aims at developing interfaces for individual users; however it is anticipated that collaborative application of these tools will be considered.

3.1.1.5 Kind of information

Types of history information available

As discussed above, the content of history information to capture is highly dependent on the user's goals in using this information. Many different types of history information are available in computer applications, such as who carried out the actions, what was done, why something was done and how. When and where are also important in information-seeking contexts, temporal order of events, and source limitations in information seeking are important aspects. One of the prerequisites of developing usable search history displays is the development of a suitable data structure.

Visibility of purpose

The visibility of purpose reflects whether the *why*-s of actions are captured and represented or not. Often, the fact that something happened is not enough information for the user, why something happened may add crucial information to the facts. Purpose information will often have to be added by the user and cannot be captured passively. The *why* of actions is separated out from the other history information types as it often gives meaning to history information.

3.1.1.6 Proxemic versus distemic

Proxemity (and distemity) are used in urban planning and social anthropology to describe the closeness of people and spaces. In relation to interaction history systems, proxemic systems are perceived as transparent to users, they are comfortable to use, easy to understand and familiar. An example of this would be a history system where the user has close control over the history information, e.g., a web-browser history list. A distemic system on the other hand appears to be opaque to users, it is an environment that is hard to understand and operate in. Users have no control over history information and the meaning of this information may be unknown to them.

3.2 Search-specific dimensions

In addition to the generic interactions history system dimensions, information system-specific attributes are defined in the following sections for the purposes of this dissertation. These characteristics define important features of the search history scope and the information system context influencing the use and implementation of search history features. These dimensions are presented in two groups: the scope of search histories (Appendix A: Framework 1) describes the various scales which can be used to define what is included in a history. The context (Appendix A: Framework 2) section defines external attributes that have an impact on how search histories are used.

3.2.1 Scope

The scope of search histories refers to how widely user and search system actions are considered when selecting them for saving. Search history information can be defined along several dimensions, such as time or types of actions, the next sections describe these and define where the dissertation research draws boundaries to search history information.

3.2.1.1 Time span

Time span describes the length of time a search history should cover, the time it should be available for. The steps of the current search are recorded, how long should they be kept and provided to the user? Just in the current session? Or for several sessions on, maybe a few months? Or should the searcher's life history of searching be preserved, stored, and provided. A related issue concerns the rules for grouping and aggregating search histories, whether a history of a single session, problem, or task is saved or whether histories arch across sessions or tasks when presented to the searcher. History information will have a different role in single versus multiple-session information-seeking situations.

3.2.1.2 Task span

The task scope of a search history refers to the domain of user activities and tasks to be included in the history. At a minimal level, queries can be saved with the number of results associated with them, but not other search or browse steps. Other steps, such as viewing or reading documents, marking or copying can be included in histories to extend the task scope to all information-seeking and some information-use activities. Recording all of the information-seeking and information-use activities is the next step, while recording all user tasks in various application leads to the area of task integration in across various computer applications.

3.2.1.3 System span: Individual versus multiple sources/systems

Related to the previous point is the question single versus multiple application environments, in the case of this research, single or multiple information sources. A single environment assures consistent object and action models, while multiple environments will warrant some kind of standardization of history information recorded across systems, databases and interfaces. The current research uses a single system with multiple databases. This context assures some standardization across databases, but also present challenges in the differences between types of documents.

3.2.1.4 User span

Search history of an individual user can often be provided to others or combined with other searchers' histories to create new information. Whether the search history only pertains to one single searcher, or is saved in order to be shared or combined is an important question for search history systems. If search histories will be shared or combined, appropriate functionality has to be provided through the user interface.

3.2.1.5 Data source: Automatically vs. manually recorded information

Automatically recorded information

Search history information can be automatically captured in computerized environments. The implications of this are enormous. In theory, anything you ever do on a computer can be recorded, replayed, reused and manipulated. This creates many legal, privacy and security issues that must be carefully considered in working with history information. Automatically recording history information also raises the question of what to record. Recording 'everything' is very expensive and in most scenarios is not necessary.

Wexelblat defines interaction history of objects as "the accumulated records of what has been done, with emphasis on sequences of actions, relationships of elements on which the user acted, and the resulting structures." This definition highlights the importance of time, relationships, and structures in histories, and thus points to elements of information-seeking interactions to record in history information. The time of actions tells us about sequence, but may also represent importance of actions and objects. Relationships of elements inform users about the whole of the context and may educate users about the system. Resulting structure can refer to knowledge structures being built, in the case of information-seeking histories. Annotated histories can be used as visualization tools for mental models of a topical area that users continuously build during information seeking.

This conceptual description of data structures will be augmented with an entity-relationship model of history information in the West legal information system context. Visualization displays build on the underlying structure and characteristics of the data and user needs. A clear structure for the recorded history data will benefit visualization techniques in many ways. It provides a prototyping environment, sets limits and boundaries for visualizations, and helps clarify the possible venues for the interfaces.

Manually recorded information. User annotations

In addition to automatically recorded information, users can contribute annotations, comments, ratings, and other information to information-seeking actions and results when executing and evaluating these. This information will later form a very valuable part of history information, as they personalize the history for the user. Unfortunately, it takes time and effort to enter this additional information, which may be a prohibitive factor for many users. Computer interfaces must facilitate user annotation of search histories. If the bar is lowered, domain expert users with complex information problems may be motivated to use these capabilities in the hopes of high returns in later stages of their work.

3.2.1.6 Dissertation scope summary

Table 3 describes the potential scope of search histories and the areas the dissertation research selected as its focus. Both automatically and manually recorded information are examined in the dissertation. Along the “Task span” continuum, search and information use activities are examined, although it is acknowledged that task integration beyond these two is an important future research direction. On the temporal dimension, current session, several sessions back and future planning are examined. Here, again, lifetime histories are promising future research avenues. Search histories can support collaborations among team members, however this study focuses on individual histories and individual use of search histories. Collaboration aspects emerged all through data collection and analysis and are discussed in the results section, but were not part of the research questions. Although only one system was used by participants, the Westlaw system has nearly 15,000 databases that searchers may search individually and integrate results from multiple sources, thus some of the issues facing search history keeping across various sources are discussed.

Time Span					
Task Span		Current session	Several sessions	Lifetime	Future, planning
	Search	X	X		X
	Search & Info Use	X	X		X
	All Tasks				
User Span: Individual user/multiple users and collaboration					
System/Source Span: Individual/multiple sources					
Data source: automatically/manually recorded information					

Table 3. Scope of dissertation research.

3.2.2 Context of search history

External factors, such as the organization, the searcher, and the task heavily influence the uses of search histories, these are reviewed in the next sections. The dissertation research focuses on domain expert searchers who have some level of search expertise, lower levels for attorneys and higher levels for law librarians. The tasks selected for the research were complex, topically-driven searches, as the assumption was made that search histories would be most beneficial for these tasks.

3.2.2.1 Domain

The domain of the search system and the search history functionality creates many opportunities and limitations. The specific content, user, and search technique characteristics of the legal information domain are described in section 2.3, along with implications for search history use.

3.2.2.2 Organizational context

The organizational context and the culture of the organization influence information-seeking behavior through accepted practices, training, systems available, and other contextual factors. These can influence the use and usefulness of search histories for the individual searcher. If reporting of time spent on tasks is important, the administrative uses of search histories will be emphasized, while if teamwork is an important requirement, sharing search histories will become crucial.

3.2.2.3 User characteristics

Information system designers and information intermediaries have long known and emphasized the importance of user characteristics in different information intermediation processes. Domain, system, and searching knowledge define user types and have effect on users' information- seeking skills and techniques.

3.2.2.3.1 Domain expertise

People with high levels of domain knowledge in the field of the information needs are capable of solving problems in the area quickly and effectively (Marchionini 1995) They understand the field and see the overall structures. They interpret results while searching and make relevance decisions that influence their consequent actions.

3.2.2.3.2 Searching expertise

Searching expertise refers to the knowledge and use of searching tools and techniques, such as using different sources and relating these to each other, the use of Boolean operators, fielded searches in automated systems, etc. It involves an understanding of the workings of information retrieval systems, computerized and paper-based, including, but not restricted to an understanding of the relationship of actions and outcomes. Searching expertise transfers from system to system, as it concerns the deeper conceptual structure of information seeking tasks.

3.2.2.3.3 System expertise

System expertise, on the other hand, relates to a specific information source or environment and involves the mechanics, the syntax of actions in these systems. This knowledge does not transfer between systems.

3.2.2.4 Complexity of information-seeking task

The complexity of information-seeking tasks varies from known-item searches to multiple-step topic-driven explorations or searches for complex information. Tasks and goals involved in these two contexts can be very different from each other. This research focuses on complex tasks, as history information can be most beneficial in these contexts. Both memory augmentation and navigation tasks are more complex in these tasks. While a user may want to go back to the results of a known-item search or re-execute a search, these situations pose less complexity.

3.2.2.5 Access variable

Access variables, such as pricing structures, speed, geographic availability play important roles in users' search history use practices. Accounting for these in the specific system context is very important. One of the most important access variables is the particular system and interface used. An information system can have separate interface with varying functionality that can influence the use of search histories.

3.3 Research questions

This dissertation set out to explore searchers' use of search history in information seeking and to design displays supporting users' tasks. The overall goal of the research was to develop a theoretical framework to understand and explain search histories and their role in information seeking, and base user interface tools on this framework. This overall goal was operationalized in three subgoals, all of which fed back into the development of the framework. These were the following:

- to examine and understand user information-seeking behavior with special focus on search histories, and to begin to develop a theoretical framework to describe this;
- to generate design recommendations for computer interfaces of information systems based on this framework and create design prototypes to illustrate and support the framework and the design recommendations stemming from it;
- to evaluate the prototypes to validate and iterate the framework and the design recommendations.

To achieve these goals, the following research questions were defined:

(1) What user tasks can search history information support and how?

For what purposes do searchers use recorded and remembered search history information?

How do they use search history information?

What kind of management tools do searchers need to use search histories?

Discussion: Chapter 5: Findings: Search history use by human user

Framework facets: 3. Search history use

4. Search history and results management

(2) What search history information to record?

What search history information is important for legal information seekers?

What level of detail and granularity of data should be captured?

What search history information do searchers remember, and record externally?

Discussion: Chapter 6: Findings: Search history data

Framework facets: 5. Search history data

(3) What user interface tools and functions are need to allow the user to use the search history information in support of his tasks?

What explicit history displays can be used to support searchers?

How can history information be integrated with other information-seeking displays to enhance their informativeness?

How can the system use search history information in supporting the user? (help, advice)

Discussion: Chapter 7: Findings: Interface design

Framework facets: 6. Interface design

The outcomes of the dissertation include the search history framework and interface design recommendations. The framework was introduced and the first two facets (Scope, Context) described in this chapter. The remaining four facets are described in Chapters 5, 6, and 7. The results are organized into these chapters to reflect these research questions above, as follows:

Chapters 5 and 6 describe current practice complemented with implications for future interface designs and Chapter 7 elaborates on new designs necessary to support legal information seeking and work.

4 Methodology

4.0 Introduction

Since this dissertation explored a new area of searcher behavior not examined thoroughly previously, qualitative methodology was selected. As described in section 4.4, qualitative methodology is especially well-suited for exploring new areas of research, as it allows for the examination unknown factors and relationships without having to define these in advance. In this dissertation, data collection was informed by earlier theories of cognitive and information-seeking behavior, but many of the facets involved were undefined at the outset of the research. Qualitative methods are often used to examine human information-seeking behavior. (Kuhlthau 2001, Barreau 1997, Marchionini 1995), and in particular, history keeping methods have been explored using these methods in Spink and Goodrum. This chapter describes the methods used in detail, including various methodological considerations.

Sections 4.1-4.4 discuss important general issues of the methodology used. These are:

(1) spiral development,

(2) the role of the system in the methodology,

(3) the role of the users in the methodology,

(4) qualitative research methods, including data analysis principles, and the data analysis software used.

Sections 4.5 and 4.6 describe the two phases of the research:

(5) Phase 1: data collection through literature review, reference librarian interviews, analysis of Westlaw usability videos, observations of and interviews with attorneys and law librarians (Appendices B- E)

(6) Phase 2: iterative participatory interface design and evaluation sessions with attorneys and law librarians (Appendices F-G)

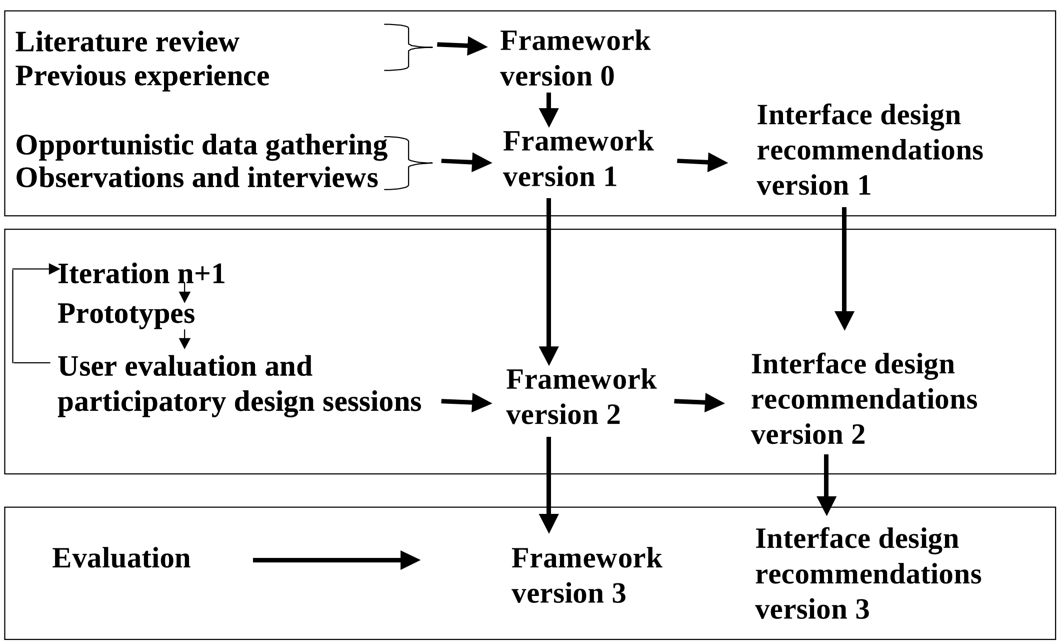
A third, future phase of the research is described briefly here and in more detail in the future research section. In this phase, interface designs proposed in the dissertation are prototyped and tested through quantitative and qualitative methods.

Section 4.7 discusses the validity and the generalizability of the study.

4.1 Spiral development

The dissertation research set out to develop a framework for understanding and describing search histories and their use in information seeking, and to design user interfaces based on this understanding. A spiral, iterative methodology was used to develop this framework. The process circled through two main iterations in Phase 1 and Phase 2 with input from different sources. Framework development was carried out in parallel with these iterations, and interacted with them. Version 0 of the framework was developed based on the literature review, and previous experience. Version 1 of the framework incorporated input from the analysis of the data collected in opportunistic data collections: reference librarians interviews and an analysis of West Group usability testing videos. The framework was used to develop a methodology for further observations of and interviews with attorneys and law librarians. After completing the observations and interviews, the framework was updated with the new findings, and a requirements statement and interface design recommendations were prepared. These designs were used in the Phase 2 participatory interface design and evaluation sessions. These sessions resulted in more interface designs and input to the framework describing the characteristics and use of search history information in information seeking. This helped update the framework to the final form presented in Chapter 3 and Appendix A.

Figure 3. Iterative framework development processes and products.



The iterative methodology was selected to avoid some of the pitfalls of linear processes. The traditional waterfall approach to system design progresses from system analysis and requirements specification to design, prototyping, implementation, and finally testing. It is called the waterfall model because the work flows from one step to the next. (Hix & Hartson, 1993) This approach has proven to have many problems and was replaced by iterative, spiral approaches in software development and interface design. The three iterations described in this research can be seen as an application of iterative rapid prototyping, although the methodology was adapted to suit the needs of this research.

Connell and Shafer (1995) define rapid prototyping as “user-defined dynamic requirements modeling” (p. 14) The context of this definition is software development and the authors differentiate conventional requirements modeling from the dynamic process. In the conventional waterfall process the users sign off on the requirements statement before development begins, while in the dynamic process an operational software model is developed and demonstrated to users in order to expand the requirements statement and verify the implemented functions. The prototype development is based on rapidly created specifications. Prototype iteration repetitively modifies the specifications; the software model, the specifications, and the prototype together form the requirements statement that users sign off, on and performance testing and tuning leverages user-approved prototypes into full-scale operational systems. This definition places the emphasis on the prototyping process in commercial software development environments, but it is useful as well in illustrating the iterative nature of the process and the importance of needs/requirements analysis in prototype development and vice versa.

In the process described in this dissertation, theoretical framework development was added to the process. It guided the collection of needs assessment data and helped the development of interfaces, and was in turn enriched and validated by those.

In the context of the current research, in Phase 1 qualitative research methods were applied to develop the theoretical framework describing the role of search history in information seeking through three different data collection efforts: reference librarian interviews, analysis of West Group usability testing videos, attorney observations and interviews. In Phase 2, user interfaces were developed through iterative user-centered interface design and prototyping methods.

These two different sets of processes are presented in a linear fashion here for clarity, but did overlap and develop interactively. In other words, the user needs assessment and theoretical framework development did not stop at the end of the first iteration, but continued and were enriched during the design and evaluation of the interfaces. The design of interfaces, although described as the second set of methods, had started from the very beginning of the research as a tool for applying and interpreting the theory and as a tool to stimulate discussion for the clarification of tasks and needs in the user needs assessment.

4.2 System/interface development

Rapid prototyping tools were used in developing the interfaces. The first interfaces were hand-drawn on paper and demonstrated to experts for feedback. Others were created by cutting and pasting printed screen images, or combining them with hand drawing. Improved versions of these paper prototypes, and preliminary computerized prototypes were developed using Hypertext Markup Language (HTML) and Visual Basic, though these were not functional prototypes. These designs were used in Phase 2 to stimulate ideas and help users envision tools and capabilities and thus elicit need- and task-statements from users, and to help the researcher formulate ideas and contextualize theoretical findings to the problem at hand.

4.3 User involvement

The methodology is user-centered and aims at involving the users both in identifying ways to use search histories and in the design of interfaces. Phase 1 used qualitative, exploratory methodologies to learn about search history use as seen in the users' own experience. Phase 2 used participatory design emphasizing the importance of involving users in the design process. Thus in Phase 2, users' involvement was very intensive; users participated in the project as co-designers. Their ideas and input were integrated into the different iterations of the prototype.

4.4 Qualitative methodology

Qualitative methodology was selected in this research, as the goal was to explore a new area of searcher behavior not examined thoroughly previously. Qualitative methods are suitable for discovering qualities, characteristics, and processes of human behavior and thinking, which are the heart of the research. Maxwell describes two of the strengths of qualitative research as identifying unanticipated phenomena in exploratory studies and focusing on the process instead of the outcome. Exploratory studies can form the basis of new grounded theories drawn from the data and help researchers identify variables for quantitative studies later. The use of search histories is embedded in the process of information seeking, as experienced by the individual user. Focusing on the process rather than the outcome enables the researcher to explore the use of continuously recorded search histories through multiple interlocking steps of finding and using information – thus looking at how this information can help searcher find information and integrate the information-seeking and information-using processes.

4.4.1 Data analysis software

The NVivo software was used for data analysis all through the research. This software package supports the management and analysis of qualitative data, such as notes, transcripts, and documents. It enables researchers to develop a coding scheme, and code any text related to the project by these codes. After the coding is completed, the software creates reports and figures from the coded text to help researchers analyze the data and recognize trends. The creators of the software package suggest the input of not only interview transcripts and observation notes, but also any other documents, such as researcher notes, proposals, etc. Thus, the management of all the documents related to a research is facilitated.

4.4.2 Data analysis

The data analysis methods used were similar in the two phases of the research project, although slightly modified to accommodate the different data collection methods. Specific methods used are described in the section of the individual data collection.

Miles and Huberman (1994) suggest early analysis of qualitative data. They propose that researchers start analyzing their data while collecting it for two reasons. First, it makes it possible to collect data to fill in gaps, and also makes it possible to create new hypotheses from the early analysis of the data, thus enabling the collection of data to address them. It can improve the data collection process while it is going on. Second, it makes the task more manageable by distributing it over a longer period. It also permits the creation of interim reports in a project. This principle was followed in this project, and data were analyzed from the expert searcher interviews, the usability videos, and the two pilot domain expert sessions before devising the protocol for data collection for domain expert user interviews and observation.

The search history framework was used as the coding scheme for coding think-aloud protocols and interviews and the observation notes with two additional facets:

(7) Implementation stage of prototype used in Phase 2

(8) Source of design idea (participant or researcher)

In Phase 1, the transcripts of the interviews and observations (except for the reference librarian interviews and usability testing videos, analysis procedures for these are described under the specific headings for the data collection) were coded at the sentence level. After interviews were coded by the researcher, the coding was examined by Dagobert Soergel, and changes to the coding were made based on his suggestions. When all transcripts were coded, reports were created for each facet, containing all the transcript segments coded under the various nodes of the facet. The first step of the analysis created a list of issues discussed in the facet, which was usually different from the list of nodes in the facet. Quotes from the transcripts were linked to the issues described. The final write-up was based on further processing and reorganization of these reports. Throughout this process changes and additions to the framework emerged.

In Phase 2, notes were taken from the audio and video recordings and they were transferred to interface design reports. Available resources did not allow full coding of the Phase 2 materials. Since there appeared to be no new issues of search history use and data needed, only interface features were coded.

4.5 Phase 1: Developing the conceptual/theoretical framework, user needs assessment

Goals:

(1) describe the role and use of search history information in information seeking in a theoretical framework

(2) collect requirements for search history displays

The user needs assessment built on different sources of information to develop a conceptual framework of how searchers use search history information and to outline tasks and requirements of search history displays. Collecting user needs assessment data from multiple sources is very important in a thorough investigation of human behavior and the definition of the requirements for design. Validity of results can be achieved through triangulation of data sources and examining data sets from various settings. The sources and methods of data collection were the following:

Phase 1:

(1) Researcher's previous experience from interface design projects

(2) Analysis of existing published knowledge (literature review):

Cognitive foundations (Section 2.1);

Information-seeking and the role of search histories (Section 2.2);

Legal information seeking (Section 2.3);

Interaction and search history systems (Section 2.4).

(3) Analysis of usability testing videos of a legal information system interface (Section 5.1.1).

(4) Interviews with reference librarians in a special library about their practices of recording search history information and techniques (Section 5.1.1).

(5) Observation of and interviews with attorneys working on complex information-seeking tasks in computerized legal information systems (Chapter 5-7).

(6) Interviews with law librarians about their practices (Halvorson, 2000) (Chapter 5-7).

4.5.1 Previous experience of the researcher

Previous personal experience of the researcher highlighted the need for search history information in information-seeking interfaces. Both individual experience as a searcher and collective experience from several information-seeking interface design projects have pointed in this direction. The researcher's personal experience and insight is an important part of the inquiry and often critical for the understanding of the phenomenon, and is weaved in through the dissertation. Collective experience from several interface studies at the Human-Computer Interaction Laboratory and the College of Information Studies, both University of Maryland, have also revealed the need for research into the use of search histories in information seeking.

The two basic validity threats in qualitative research described by Maxwell (1996) are researcher bias and reactivity. Bias is based on the researcher's theories and preconceptions formed before data collection and can result in the selection of data that enforces these preconceived ideas. The researcher in this dissertation brought to the problem an understanding of basic information-seeking behavior phenomena and familiarity with the interface design process. The first three data collection activities (literature review, reference librarian interviews, usability testing video analysis) set out to explore whether search histories are used in searching or not. After the analysis of both data set verified the need for search history information, the research question became an exploration of what information need to be recorded, what user tasks it can support and what user interface tools are needed to manage and build on this information. The researcher had some assumptions based on the previous data collection, but was open to explore user tasks that used search histories and tasks that did not require a history record.

4.5.2 Literature review

The conceptual context of the research embodied in published literature is an important part of the theoretical framework. The results of the literature review had informed the research before the main data collection steps started. The results of the literature review were discussed in Chapter 2. They describe cognitive factors relevant to the use of memory and search histories in information seeking, information-seeking processes and steps with special attention to phenomena where search histories can be helpful. Legal information seeking and legal information users are also reviewed, and finally interface design techniques for search and other activity history systems are described. All four sections of the review focus on implications for search history use, and the results of literature review, along with the researcher's personal experience, was used to form Version 0 of the framework, which in turn provided the basis for developing methodology of the data collection in Phase 1.

4.5.3 Reference librarian interviews

The researcher had the opportunity to interview reference librarians in a special library about their personal methods of archiving requests they received, and searches they carried out in response to those. These interviews were carried out in the framework of the implementation and customization of a request-tracking system that the special library was adopting. The interviews served as the user needs assessment for the customization of the request-tracking system and also provided another input to the motivation for the dissertation research.

Participants

Five reference librarians in a busy special library setting were interviewed. All five of the librarians had several years of experience in reference work. Four of them worked in the same library for several years, one of them just transferred from a smaller library in a partner institution. In addition, five other experts from other parts of the organization were interviewed who had previous experience with using the automated request tracking system to be implemented in the special library.

Materials

A flexible interview guide addressing librarians' history procedures. The guide was flexibly adjusted to each individual interviewee.

Procedures

The interviews were guided by a flexible interview guide, but were modified in each case to suit the individual interviewee. Participants were interviewed in their own offices, and were asked to demonstrate how they carried out their task, with special emphasis on the recording of search information. This choice of setting enabled the researcher to review examples of physical external representations of searches that users created. The researcher took notes of the interviews recording comments librarians made and also describing physical artifacts used by librarians. Examples of these were also collected. The interviews were audio taped, except for two participants who did not agree. The tapes were transcribed, summaries of the notes and the transcripts were sent back to librarians for member check, and a final report was created after confirming their content.

Analysis

The researcher transcribed her notes in formal reports right after each interview. These transcripts were then forwarded to the reference librarian interviewed (member check) and were updated with the feedback of the librarians. The individual reports were summarized for the purposes of the request-tracking system customization.

For the purpose of the current research, the interview transcripts were coded, thus providing the first version of the coding scheme later used in coding the user interviews and observations. They were revisited and the coding updated after the coding scheme was finalized. The findings from these interviews also influenced the creation of the interview guide for the larger study. The summary of the results can be found in Section 5.1.1.

4.5.4 Usability testing video analysis

The West Group has carried out usability testing of its different software interfaces. Seven of the usability test videos were selected for inclusion in this phase for their coverage of search history tools. The researcher was authorized to watch the video tapes on site and take notes from them, however was not allowed to copy the tapes or create transcripts. These sessions have been examined with special focus on tasks and questions related to search history functions of the software. The findings served as motivation and guidance for the later observation and interviews, and are summarized in Section 5.1.1.

Participants recorded in the videos used two different software interfaces of the West Group legal products in the usability testing sessions. The interviewer gave the participants a series of tasks and asked them questions about their reactions and use of the interface. Special interest was given to the "Project" feature in the interface. This feature allows users to record the steps of their search and save the record in a project. Participants were asked about the "Save search," "Project," "My Favorites," and "Search Trail" features of the interfaces; all of these support some kind of search history mechanisms.

Participants

The participants were intermediate and expert users of West Group software interfaces, recruited and interviewed in different parts of the country. The length of the sessions was about an hour, but only selected sections were reviewed for the purposes of this research.

Materials and procedures

The task protocol and interview guide of the usability testing sessions drove the data collection. Only those questions were examined that referred to search history-related functions in the interfaces. Locating these video segments was a time-consuming process, as the sessions were of differing length and pace. Irrelevant portions of the video tapes were not viewed. Notes were taken about the reactions of the participants to the questions concerning search history tools in the interfaces.

Analysis

Notes taken at the West Group site were analyzed by grouping issues about the use of search history tools. Results are summarized in section 5.1.1.

4.5.5 Think-aloud and observation of attorneys searching with follow-up interviews

Participants

Two pilot participants were followed by six other participants in this part of the study. Searchers were recruited from among attorneys and law librarians in the Washington, DC metropolitan area. They were selected based on their legal domain and legal searching experience; professionals who searched for information as part of their job and were familiar with electronic legal resources were selected. The participants represented somewhat differing levels of domain and searching expertise. Law librarians are trained in searching, have experience in using search strategies and techniques, and are most likely more familiar with the Westlaw interface than attorneys. The pilot participants were attorneys who were studying at the College of Information Studies and thus had some amount of formal searching training. The six participants were all attorneys up to their sixth-seventh year of practice when they still performed information-seeking tasks as part of the job descriptions. Attorneys are also trained in searching, but not with as much detail and emphasis as librarians. At many law firms, attorneys must participate in training sessions for the Westlaw system in order to be able to use the service.

Current and former attorneys, four males and four females with a mean age of 38 years, volunteered to participate. Participants had an average 11 years work experience working in the legal profession, and were frequent searchers familiar with advanced search strategies and legal search systems. Half of the subjects are current users of the Westlaw system discussed in this study. Further details of the demographic data is presented in Table 4.

Time per week spent searching	<4 hour: 3 4 to 10 hours: 2 > 10 hours: 3
Comfort with Boolean operators (1=Very comfortable, 9=Not comfortable)	3.4
Comfort with West Key number system (1=Very comfortable, 9=Not comfortable)	6.0
Comfort with Natural language searching (1=Very comfortable, 9=Not comfortable)	3.9
Have you taken a legal research course	Yes: 8 of 8
Length of time using electronic info resources	< 6 months: 1 >5 years: 6 NA: 1
Length of time per week on Westlaw	<1 hour: 6 1 to 4 hours: 1 NA: 1
How many computer applications have you worked with	< 12: 3 > 12: 5

Table 4. Demographic data of attorney observations and interviews participants.

When participants from this data set are quoted, the following designations are used:

Pilot participant 1: P1

Pilot participant 2: P2

Participant (searcher) 1: S1

Participant (searcher) 2: S2

Participant (searcher) 3: S3

Participant (searcher) 4: S4

Participant (searcher) 5: S5

Participant (searcher) 6: S6

Materials and procedures

The Westlaw legal database system was used at westlaw.com. This system is widely used in legal contexts, most law students have free access to it during their studies and is used in many law libraries and law firms.

An informed consent form was presented to the participants to inform them about the conditions of the study and to elicit their agreement. (Appendix E) A demographic questionnaire was administered to the subjects in the first part of the interview to learn about their legal domain, Westlaw system, general computer, and searching expertise and experience. After filling out these two forms, participants were introduced to their tasks. During the actual search sessions participants were asked to think aloud and the researcher prompted them to comply with this request. This was necessary to elicit any problems and issues where search histories would have been useful. While the participants were searching, the researcher observed their actions and took notes of the process. The search sessions took place in the participants' own offices.

Participants were asked to search the Westlaw databases for about an hour, although the actual length of the searching section of the protocol varied from about an hour to about three hours, depending on how long the attorney wished to search. The sessions were audio-taped. They searched their own topics of interest, or if none of those are available, they were provided with complex, topical searching tasks.

Participants were asked to introduce their searching task in the context of their work tasks. Taping of the session started at this point.

After the search session, the participants were asked first to summarize what they had just done and what they remembered from the session. Then they were interviewed based on the interview guide in Appendix B. This guide was developed on the basis of the data collected from the literature review and the initial data collection and analysis. This, along with the transcripts of the interviews and the researcher's notes formed the basis of the data analysis. As observations and interview techniques were used for data collection, these are reviewed here.

Observation

Searchers were observed while searching for information. They were asked to think aloud while searching, every instance of referring back to previous information was notes in the transcripts and coded under different nodes describing:

(1) the type of data mentioned

(2) the type of action carried out with or on the data

(3) the user task the retrieved data supported

(4) potential user interface design need

The researcher observed and recorded user actions that were not described through the think-aloud method. Searchers often create external representations of their search histories outside of the search system: in word processing files, paper notes, printout. They may refer to books or other print documents, ask colleagues or examine computer files not related to the search system. These activities may all point to the use of search histories, and are very important for the study.

As mentioned above, participants were observed in their own offices.

Studying searchers in their own environment was important in this study as their use of history keeping often extended beyond the search system to their physical environment. This can be observed more fully in their own context. Maxwell (1996) states that the observer is present or not, than on the environment and context. Thus changing the reactivity (participant reactions to the presence and activities of the researcher) is much less dependent on whether context and moving participants to the researcher's site would have resulted in stronger reactions from the participant than the researcher being present in their office.

One of the factors that differentiate observation methods is the degree of involvement of the observer. The current research examined individual behavior on the part of the subjects, thus the observer remained mostly distanced from the activities. She did, however, answer questions related to the system and discussed some aspects of searching, but not to the extent of making her a participant in the activity.

Interviews

Patton (1987) states that the goal of interviews is to find out from people that which cannot be observed. Their perspectives, opinions, and feelings cannot be examined through observations; interviews are needed to complement observation data. . Maxwell (1996) adds to this definition by including process and event data to be gained from interviews. This is especially true of events and processes that cannot be observed for some reason, for example past events. Interviews can be most helpful to elicit descriptions of events or processes if the interview questions address specific events instead of general questions.

Interviews followed an interview guide. The exact wording or order of the questions was adapted to the individual participant because participants followed different patterns of searching and because they recalled and analyzed specific events, making it impossible to anticipate the exact order and wording of the questions. Questions often referred to incidents that came from observations or other field data; the searchers were asked to explain or react to critical events that the interviewer recalled or described.

The interview after searchers finished searching explored their search history use further. The following questions were posed to the searchers:

(1) Can you summarize the searching you just did?

(2) Please rate how important saving these different items is for different time periods. (Appendix D)

(3) Please look at the top five items, why are these so important? How would you use them?

(4) What other techniques do you use to help you remember information?

  (4.1) What do you do that helps you remember your searches and results now?

(4.2) How do you create and organize reminders? Can you show me some examples?

(4.3) Do you use this information again? If yes, how?

(5) Do you use history mechanisms in online searching when those are available in the system? Why? Why not?

(6) Do you ever save your queries? Why? Why not?

(7) Do you save you search results? Why? Why not?

(8) Do you use the project feature in the Westlaw PC client interface? Why? Why not?

(9) Do you use bookmarks or the ‘Go’ list in the web interface? Do you use search trails in the new web interface? Why? Why not?

Answers to question number 1 were analyzed to find out what information searchers assumed to be important enough to remember. Questions 2 and 3 aimed to measure the perceived importance of recording various history items for various time periods. Questions 4-9 explored other areas of search history use, they were posed flexibly, depending on the searcher’s experience and topics covered earlier in the session. In addition to the formal interview questions, more informal, follow-up questions were posed to participants during the interview to explore topics that emerged during the conversation or user actions observed earlier in the session.

Data analysis

Demographic questionnaires, taped think-aloud data and interviews were transcribed after the sessions and entered into the NVivo software. The data analysis was carried out as described in Section 4.0.6.

Results of Phase 1

The results of Phase 1 integrated findings from these different sources. They were embodied in the first version of the theoretical framework and requirements statement, and initial prototype designs based on these. The framework included a description of how searchers use search histories currently, and potential new ways search histories might improve searching. The requirements statement provided a list of areas of information seeking where search histories may be useful in terms of the user interface. This iteration also included a set of preliminary interfaces showing some of the ideas from this iteration.

4.5.6 Published interviews with law librarians

In addition to the various data collected by the researcher, transcripts of interviews with law librarians were analyzed to inform the research. These interviews were conducted by T. R. Halvorson who sought out and spoke with eight outstanding legal research experts about their practices in searching for legal information and in educating others about legal research. T. R. Halvorson authorized the use of these interviews for the current research and provided electronic versions of the text of the interviews published in his book (Halvorson 2000) in the Super Searcher series. The goal of the series in the words of the editor, Reva Basch, is the following:

> "The Super Searchers series is an attempt to convey, in their own words, a sense of how expert researchers think: how they conceptualize, plan, and carry out a research project; why they take one approach rather than another; what prompts them to select a particular source or medium; why basic skills and training are merely the foundation on which intuition, creativity and ingenuity all build"
>
> (Halvorson 2000, p. ix)

The author of the book used the expert-interview approach basing the questions of the interview on previous volumes in the Super Searcher series. The interview guide was flexible, and each interview followed a somewhat unique set of issues. These interviews were included for several reasons. Triangulation of the data set, examining users in various settings helps insure the validity of data. In addition, expert searchers can highlight highly effective methods and procedures for legal information seeking, thus providing a rich basis for interface design.

The interview questions were naturally different from the ones used in the sessions conducted by the researcher, but 20-50 % of the interview text was relevant to the research questions posed in this study. These interviews were analyzed using the same methods as with the transcripts from the attorney observations and interviews. When quoting participants from these interviews, T. R. Halvorson is represented as: Halvorson, the interviewees by their last names.

4.6 Phase 2: Interface designs

Goals:

(1) iteratively work with participants to develop interfaces based on the findings of the user needs assessment and the theoretical framework,

(2) update theoretical framework.

The second iteration of the research involved the further development of the user needs analysis and the theoretical framework through the design and prototyping of interfaces, and the testing of interfaces. All these activities were carried out in iterations, building on the feedback from users about the previous versions. The emphasis in this iteration was on the design of interfaces, as reflected in the naming of the iteration. The design methodology used involved users throughout the design process and proceeded through several iterations. User involvement in design does not only insure the usability of the interfaces but may add to the understanding of user behavior described in the theoretical framework thus modifying the outcomes of Phase 1.

4.6.1 Participatory design

Hix and Hartson (1993) suggest the involvement of a few carefully chosen users in the design process. Users on the design team can bring a thorough understanding of the tasks and needs of the domain that designers would not be able to amass. They may raise issues that the designer could not because of the discrepancy between designers' model of a task and the actual task (Nielsen 1993). User representatives should be selected from each class of users. The presence and contribution of users in the design process can turn users into stakeholders, thus promoting the acceptance of the interface. Although this was not necessary in the current research, user involvement enriched the design of the interfaces.

Shneiderman (1998) points out the dangers of involving too many users in the process. This would lead to excessive costs and may lengthen the design process. It can also lead to compromises based on needs of incompetent users and difficulties created by rejected ideas. The selection of users to participate in design is a very important step that needs consideration. Participants should be told about their role and commitment in advance.

Nielsen (1994) highlights the importance of presenting the prototypes to users in a form that they can understand and react to. The application of visual display instead of textual descriptions is crucial. Another important consideration is the aesthetic appeal of the prototypes that may be a crucial factor in receiving adequate feedback. A good interaction design may be hidden behind an aesthetically badly executed or malfunctioning interface if those issues distract users' attention and leave a negative image of the prototype. It is also important not to simply ask users what they want to do, because often they cannot anticipate their needs due to their lack of knowledge of technology. Nielsen cites the results of Root and Draper (1983). They found no relationship between users' ratings of features before and after trying them.

4.6.2 Iterative design

The spiral methodology used in the research was described in Section 4.2. It is important to highlight here the iterative nature of the interface designs and requirements. The first prototypes were created on paper and presented to interface experts, search experts, and domain experts. The comments from these users were integrated into the subsequent iterations of the prototypes. The development of these interfaces and requirements through the several iterations is described in Chapter 7.

Hix and Hartson (1993) mention the unpredictability of human behavior as one of the motivations for iterative design. Thus, the design of interfaces, and especially interaction "must be essentially and inherently iterative … a self-correcting process." (p. 97)

Establishing requirements for each iteration is suggested in commercial software development projects. This practice was followed in this research to better specify the iterations and the deliverables. The iteration of requirements for interfaces is closely related to the iteration of the prototype designs. The responses from users in each iteration were integrated into the requirements document.

User interface design guidelines

The literature of human-computer interaction and interface design provides many interface design principles and guidelines. These are adhered to in the design of interfaces and used in heuristic evaluation and somewhat is expert reviews. For overviews of these guidelines see Shneiderman, Hix and Hartson, etc. Some of the guidelines important for search history displays were reviewed in Chapter 2.

Participants

Four participants were involved in two pilot sessions and nine domain expert participants, attorneys and law librarians were involved in three interface design and critique sessions. Participants in the pilot sessions were interface experts, one system developer and one layperson to insure that the protocol is clear and the design tasks are described adequately for non-interface or system development experts, as the legal domain expert participants were.

Five library professionals, two attorneys, and two other legal professionals, four males and five females with a mean age of 37 years, volunteered to participate. Participants had an average 8 years work experience working in the legal profession, and were frequent searchers familiar with advanced search strategies and legal search systems. Further data on the demographic information is presented in Table 5.

Time per week spent searching	<4 hour: 1 4 to 10 hours: 1 > 10 hours: 7
Comfort with Boolean operators (1=Very comfortable, 9=Not comfortable)	1.9
Comfort with West Key number system (1=Very comfortable, 9=Not comfortable)	3.7
Comfort with Natural language searching (1=Very comfortable, 9=Not comfortable)	2.8
Have you taken a legal research course	Yes: 6 No: 3
Length of time using electronic info resources	>5 years: 9
Times per week spent using online or CD-ROM information sources	5 to 7 times: 1 > 10 times: 8

Table 5. Demographic data of participatory design participants.

Materials

A demographic questionnaire was used to elicit information about the participants. In two of the sessions an additional observer was present to take notes as the researcher was very involved with recording the design process and did not have the opportunity to take notes.

Notes from the interview and the observation, and transcripts of audiotapes of the sessions were all entered into the NVivo system for data analysis. An interview guide was used for the rest of the session. Both the questionnaire and the interview used neutral questions in order not to influence users' responses about the different features, although value judgments were added in the design section of the sessions. Since different interface features were to be tested in the different sessions, the session protocol and the questionnaire were modified for each session to reflect these differences.

Procedures

The two pilot sessions allowed the researcher to test and develop the protocol for the four sessions. After analyzing results from the two pilot sessions, three variations of the protocol were created for the three sessions:

Session 1:

(1) Introduction

(2) Discussion of search history concept

(3) Demonstrate and critique existing search history interfaces

(4) Design search history interfaces

Session 2:

(1) Introduction
(2) Discussion of search history concept
(3) Demonstrate and critique existing search history interfaces

Sessions 3 & 4:

(1) Introduction
(2) Discussion of search history concept
(3) Design search history interfaces
(4) Demonstrate and critique existing search history interfaces

The variations created ensured that neither the participants own designs, nor the designs demonstrated to them by the researcher overwhelm the creativity and critical point of view of participants and that no design dominates the discussions.

In some of the participatory design sessions, an outside observer was invited to take notes as the researcher was busy with conducting the session, and both audio and video tapes were recorded. Audio tapes were recorded in addition to video tapes to adapt to the transcription technology available to the researcher. Video tapes were required to capture visual data. In addition to video tapes, a digital camera was used to take pictures of designs. An analysis of all of these data sources was applied to ensure that information was not misinterpreted or misunderstood.

Participants were introduced to interface designs and asked to imagine using them and describing positive and negative features about them. When they were asked to design interfaces, they were introduced to some interfaces as examples and then were provided with tools to create their own designs. The tools included paper, colored pens, printed interface design elements (windows, scrollbars, buttons, and sop on), scissors and glue. Using these tool participants created their own interface designs.

Data analysis

In Phase 1 notes were taken from the audio and video recordings, the designs were reproduced in digital format and interface design recommendations created based on these. Many the comments made about search history use in Phase 1 were repeated in Phase 2 and thus reinforced.

4.7 Validity, generalizability, and limitations of study

4.7.1 Validity

How shall we know that the conclusions reached in this research are valid? Are the descriptions of search history uses accurate and complete? Are the interpretations of user actions and needs correct? Validity is the correctness or credibility of the accounts of qualitative research, such as interpretations, descriptions, and conclusions. It does not assume to be identical with an objective and unique truth, but a good standing point that can be proved valid. The main point is to identify and rule out the ways that the research may be wrong, that is to identify and deal with validity threats.

Maxwell (1996) described two basic validity threats, researcher bias and reactivity. Both of these were addressed above, in various sections of Chapter 4. Although he explains that using certain methods will not guarantee validity in research, some procedures can help rule out validity threats and test the validity of the conclusions through ensuring that evidence challenging the conclusions was sought after and taken into consideration. Member checks, triangulation, the use of rich data and comparisons are described in the context of this research.

Member check:

Whenever it was possible, transcripts of sessions were sent to the participants for confirmation that the information captured was accurate and complete. When this was not possible due to access limitations to the participants, the researcher often rephrased the participant's comments and asked for confirmation on the meanings of the comments. These checks were used so that the right descriptions of sessions and interpretations of participant remarks were made by the researcher.

The participants in the study were practicing attorneys whose time was highly priced. Access to them was limited to one session. This prevented the researcher from returning to participants and follow up, the maximum amount of information had to collected in the time allotted. In order to examine the use of search histories as true to real use as possible, participants were observed and interviewed in their own environment to ensure that they had all their reguler external tools available.

Triangulation and comparisons:

Triangulation of data collection reduces the risk of chance associations, and of systematic biases through data collection from a variety of people in a range of settings, using various methods. (Maxwell) It allows wider generalization of finding and conclusions. Comparisons are similar techniques, comparisons can be made among multiple sites of data collection, literature and new data collected and the researcher's own experience.

The current research built on several independent data collection efforts, as described above. In particular, the main body of data was co0llected from two different sets of attorneys using different sets of methods: observations and interviews, and participatory design sessions. The data collected through participatory design sessions often enforced finding from the observations and interviews, and also expanded those findings by including new aspects of the phenomena. An additional data set, the law librarian interview transcripts (Halvorson 2000) introduced another source of data that was analyzed using the same methods and provided another set of data points to compare findings with.

Use of rich data

Using rich data refers to the inclusion of detailed and as complete as possible data recording in the collection phase. Using rich data helps better identify competing alternative interpretations and information contradictory to the researcher's theories. In the current research, observations and interviews were audio taped and literal transcriptions were created in addition to the researcher's notes from the sessions.

4.7.2 Generalizability and limitations

The following questions will be described to address the generalizability and limitations of the study:

(1) Legal domain

(2) Sample size

(3) Single-session data collection

(4) Researcher's interpretation of observations and think-aloud protocols

(5) No testing of designs

(6) No data on possible negative effects of history interfaces (see Section 5.0.3)

Maxwell (1996) differentiates internal and external generalizability of qualitative research. Internal generalizability refers to the validity of research conclusions within the setting or the group studied, and in similar settings or groups. In the current dissertation, the applicability of conclusions to other legal information seeker groups and contexts would describe the internal generalizability of results. When other searching domains such as general Web searching or medical information seeking are considered, external generalizability of the research conclusions should be discussed.

This study focused on a limited task and content domain: legal information. This limitation was imposed for two main reasons:

(1) to allow the researcher to fully explore and take a holistic view of the use of search history information in a specific context;

(2) to examine an area that may include especially heavy users of search histories because of the characteristics of the domain, thus learning the most about the use of search histories.

Maxwell (1996) points out that the value of qualitative research studies often come from the lack of external generalizability, their tendency to study settings with specific forms of behavior that may not be typical of other settings but can highlight processes that are effective and exemplary or identify localized problems that are not otherwise detectable. The dissertation research was exploratory in nature and the field of legal information work was selected as one that would greatly benefit from the availability of search histories, thus offering a wide range of uses, as opposed to another task and content domain where search histories are not as important and only a limited set of uses could have been discovered. The findings are thus limited to legal information users, especially to attorneys and law librarians in large law firms usually working in teams.

Findings from this domain can also be useful for other domains, although it is clear that not all of the search history uses identified would be beneficial for other domains. Considering multiple content and task domains would have required a more superficial analysis of search history use in the various domain and would potentially have discovered a smaller set of tasks that can be supported by search histories. Further study is warranted in order task domains to find what portion of the findings would transfer well to other domains. Transfer of some results might be possible based on knowledge of how people in the other domains work. The framework developed in this dissertation can be used as a starting point in examining other domains. It provides a list of issues to examine in an organized set.

The main body of data was drawn both from end users and search intermediaries in the legal information field. Differences in searching patterns between these two groups are acknowledged and addressed in the analysis. Expert searcher techniques may be unique to that group but can suggest exemplary procedures for attorneys.

The sample size of the main data collection effort (8+8+6) was small if judged by quantitative research standards. However, sample size in qualitative research is determined by the results of the data collection. When the results of the data analysis repeat themselves over and over again, the data collection effort can be terminated. In this dissertation, both later observation and interviews and the participatory design sessions reinforced the findings of the earlier observations and interviews, similar findings reappeared and it was decided that the sample size provided sufficient information to study the research question.

Another potential criticism can be the lack of longitudinal data on search history use. Since attorneys and law librarians were only interviewed in one session, long-term use of search histories could only be examined through self- reporting in the interviews and in the think-aloud sessions. The availability of attorneys and librarians limited the scope of the data collection in this respect. The focus of this dissertation is the use of search histories in the current session. However, reports of longer-term use of search histories were frequent in the data collected, and thus results are reported.

The researcher interpreted observation and interview transcripts and coded them based on the framework. This coding was checked by the dissertation advisor to ensure reliability. Whenever it was possible, transcripts of sessions were sent to the participants for confirmation that the information captured was accurate and complete.

When this was not possible due to access limitations to the participants, the researcher often rephrased the participant's comments and asked for confirmation on the meanings of the comments during the interviews. The participants in the study were practicing attorneys whose time was highly priced. Access to them was limited to one session. This prevented the researcher from returning to participants and follow up, the maximum amount of information had to collected in the time allotted. In order to examine the use of search histories as true to real use as possible, participants were observed and interviewed in their own environment to ensure that they had all their regular external tools available. These checks were used so that the right descriptions of sessions and interpretations of participant remarks were made by the researcher. Even though these measures were taken, researcher interpretations may occasionally be wrong.

The designs created in the dissertation were not formally evaluated. The designs were presented to participants to obtain feedback, but no formal comparisons between designs and between systems with and without search histories were conducted. This is proposed for future research efforts.

Finally, data on the non-use of search histories was not specifically collected, although the topic was occasionally raised during the interviews. The results of these comments are summarized in section 5.0.3. A more extensive examination of when search histories are not helping but hindering is necessary in the case of system designs. Whether or not to implement search history features in actual systems depends on the information-seeking context of the system.

5 Findings: Search history use by human user

Chapters 5 and 6 discuss the results of data analysis from Phase 1. Many of the findings of Phase 1 were reinforced in Phase 2 data collection; however, that data was not coded in detail and thus all quotes are from the Phase 1 data set. Chapter five describes information-seeking and information-use tasks that can be supported by search history information. The tasks are described in terms of what searchers do now, what needs they expressed, and what needs the researcher identified based on difficulties participants experienced in finding and using information. Chapter 6 describes the types of search history information to record to serve these needs. The data elements are described in the context of their use. These chapters weave together specific findings on what study participants did or said they wanted to do, ideas on search history functions that could solve problems that were observed or that were mentioned by participants, and ideas that were generally inspired by the data, the literature, and analysis.

5.0 Introduction

Chapter 5 introduces ways that search histories can help users find and use information. It describes user tasks and history applications related to the following six task areas presented in Figure 4:

(1) Memory support (Memory aid)

(2) Search system use

(3) Find information/Search

(4) Search task management (supporting Find information/Search)

(5) Search history and search result management

(6) Information use

(7) Integrate with other tasks

(8) Integrate with other people

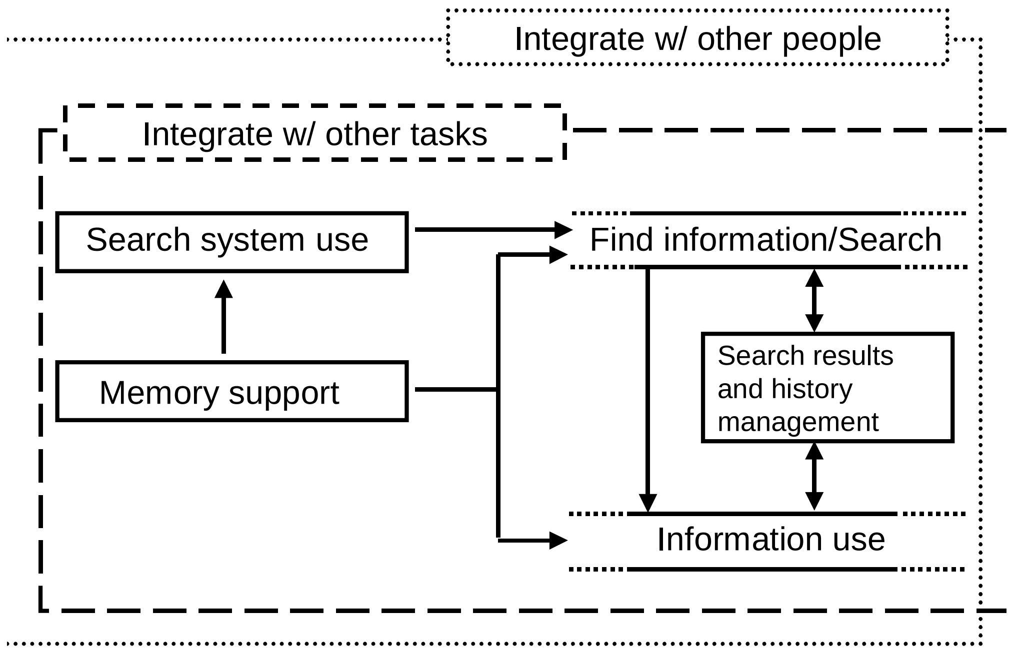


Figure 4. Search history use areas.

Memory support is the most basic function of search histories. Activity information is recorded and provided to the user thus removing the burden of remembering all activities and items from the user. Other application areas build on memory support. In order to find information in computerized systems, searchers have to be able to use the application interface. Search histories can be helpful in this respect. Finding and using information build on both of the earlier application areas, but also introduce new history-based interface functions. Looking outside of the search task, users will use the information found in other tasks. Methods to integrate search with other user tasks can also be supported by search history information. Finally, the even larger context of information seeking involves collaboration with others, where search histories can help with transferring information between team members.

Before delving into detailed discussions of these areas in sections 5.8 through 5.8, the chapter presents in this introduction some general issues. First, section 5.0.1 presents a vignette that illustrates the major finding of this study: search histories are both used and useful. Section 5.0.2 on search maps gives a particular perspective on a global view of search history. Section 5.0.3 discusses possible disadvantages of search histories and section 5.0.4 deals with legal and ethical concerns.

5.0.1 A vignette illustrating search history use

The first data collection step was interviews with reference librarians in a special library about the search history keeping practices. The librarians reported using many different computerized databases and paper resources in their work. They had a request-tracking form that they filled out with differing levels of detail. This form recorded many elements of their search histories, mostly for management statistical reporting purposes. In addition to this, they kept personal logs and reminders of searches that were reviewed in the interviews in various formats. The interviews were carried out in preparation for the implementation of a computerized request-tracking system.

In the intermediated searching context of the special library, tracking requests through the organization was an important function. The library received requests through many channels, which made their tracking challenging. In order to take full advantage of these tracks, more information needed to be saved than a mere statement of receiving and completing requests. Intermediaries process results for use for end-user patrons.

Librarians mentioned several different methods for individual "request archiving" in the interviews:

-paper folders organized by department containing printouts of emails with additional information:

-name of file, or record of databases or internet resources searched containing the answer to the request

-date sent

-content of file

-number of records

-other comments

Reference librarians wrote down database names and pointers to results. One staff member kept a search log in the case of longer and more involved requests where several searches were necessary and answering the request took longer. This log recorded what had been done and what had been sent to the user. The librarian started recording this information after the firs three or four searches.

Another librarian described an MS Word file kept with pointers to different kinds of information that the librarian considered hard to find. The entries in this file included a keyword or expression and a short description of the resource, the entries were organized alphabetically. The entries described sources, websites, usually in the form of a topic and pointers to sources of information for that topic. This was a very easy-to-use tool, with low levels of effort required to maintain and search it. This file was also printed and taken to the reference desk when the librarian was working there.

Emails sent to users and replies to those (thanks) were kept in two separate folders organized by department of the patron by one staff member. Several staff members mentioned that it was useful to have a record of a request filled because of follow-up questions received later. Another aspect is statistics, major topics, clients, types and complexity of requests were monitored and reported to management. The staff may be asked to show examples of work when this record is also useful.

In a second data collection effort, preliminary user interviews have shown that expert searchers involved in complex searching tasks require history-keeping functions from the system. Evidence was found that if these functions are available, searchers use them. In addition, they create external memory aid through taking notes to record different aspects of the searches and results. Subjects in the usability testing of Westlaw interfaces were asked about the history functions in the current system, what they use them for, and whether they feel these tools are necessary. The overall response was positive to using these tools, although the participants requested some improvements.

Searchers in the dissertation interviews and observations were asked to summarize their searching immediately after stopping the search in order to identify which elements were the most memorable in the short term. The elements remembered most were incidents of success and failure, important strategic steps and resulting documents. The major events in the search process were recalled in their order of occurrence. Highly successful and unsuccessful steps and techniques were also remembered and evaluated when the searches were summarized, along with changes in their strategy. Keywords from queries, especially new terms learnt, were listed, along with the approximate size of the result sets returned for specific queries containing the keywords. The type of the query was also often mentioned (natural language, Boolean, key number, etc.). Participants cited important documents they found and mentioned how they influenced their search strategy.

Searchers remembered this information because they thought them important during and immediately after searching. Most of the information searchers recounted can automatically be saved, such as queries, terms, size of result sets, and so on. Searchers subjective judgments about these events can only be saved through explicit user input, tool for recording this are needed. In addition to providing important highlights of searchers for the long term, search histories can help fill in the gaps in users' memory of events and provide a tool for them to easily retrieve information that was judged less important while searching and thus not remembered.

5.0.2 Search histories and search maps

Many current information retrieval systems record queries, however they do not record other equally important steps. Recording user actions is important because it represents personal, anecdotal memories for the searcher. The data recorded on user actions, system actions and information content makes up the search history of a given search session. Automatically recorded search histories can be annotated and marked by users, thus personalizing them and preparing the information for reuse.

The term “road map” was used by multiple participants in the study to describe a personalized search history. The “road map” is an automatically recorded search history, a record of user and system actions, complemented with a temporal representation of the user’s ideas, thoughts, interpretations, and reactions to system output. A road map is a highly personalized and customized search history. The exact same series of steps can represent very different meanings for searchers, thus the plain search history record with user notes, annotations, interpretations added to it form a road map of the user’s physical, behavioral and cognitive journey through the search. A “road map” describes where the user went in the system and what she found, and also describes what the information found meant to the user and how it applied to her problem.

> S2: At least before there was no way to save your Westlaw searches. So, by printing out your search history, your terms, connectors, and everything, instead of trying to recreate the wheel, you would know. And it was a lot easier; especially I used to highlight my search history, if you can… When you go back and you want to update you can actually see like the dates of the cases, and all that stuff. So it’s a good … it’s a road map.
>
> I: So it’s sort of a map of your conceptual thinking of the issues and how they relate to each other, is that right? And then a way to retrace it?
>
> S2: Yes, and instead of you writing down every single case, every single cite, you print it. Which you didn’t used to be able to do.

In addition to using the term “road map”, a different spatial metaphor appeared, another participant described the roadmap tool as a Global Positioning System (GPS) device, something that would tell him where exactly he has been::

> S6: It’s a big document, we are going to need to know … Ok, what would have been helpful in this, if I had a sort of a GPS, something that guided me through where I was so maybe I could have put the brakes on, go back to where I was.

Although study participants used the term “road map”, this term does not accurately express the meaning they described. A road map presents a static structure of physical pathways that travelers can select as part of their route. In terms of a search system, a road map would be a system design blueprint. A personalized search history that the participants described on the other hand describes a constantly changing traversal path over the structure described in the road map. Thus, “search map” is proposed and used throughout the dissertation to describe the personalized search history described by participants, representing the searcher’s path through the search system, his navigation pattern; it includes steps, input and output, documents, and also user notes, annotations, explanations and interpretations and interpretations, reflecting the user’s view of searching. Creating this search map is additional effort; as much as possible recording search histories should be automated, with an option for the user to change, delete and annotate and organize them.

The search map reflects searchers’ mental model building; however, it is highly particularized, it is an instantiation of a mental model combined with the physical actions that led to the formation of the mental model. Further differentiation between these is discussed in the sections: memory support (5.2) and development of mental models (5.6.1.2). One of the participants alluded to this definition of mental models when describing her "road map":

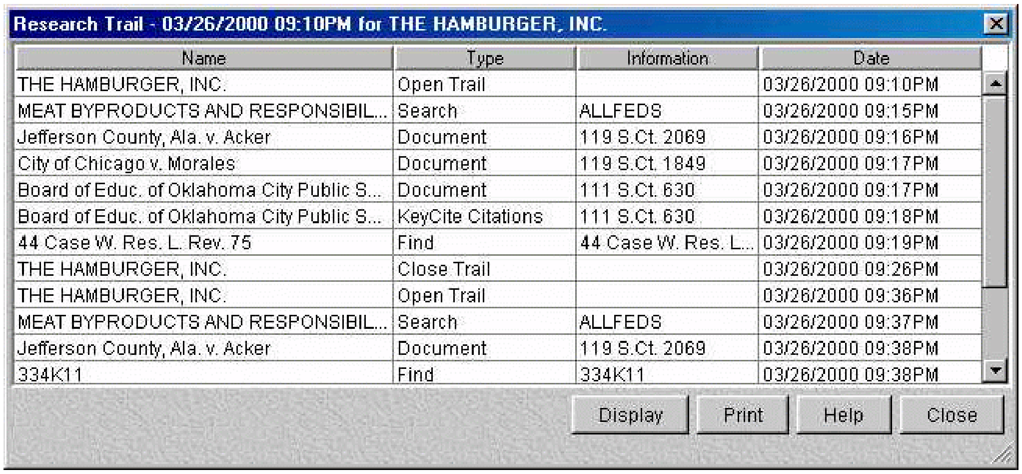
> S2: Road mapping.
>
> I: Can you elaborate on the road mapping?
>
> S2: You know I don't really know, except for an issue comes before you, how you think about the issue, how you turn it over and manipulate it, can influence your search results. You need to make sure that you are manipulating the issue in the right way. A lot of times you can see, as a young associate or junior associate, you go to the partner and they'll say why did you pull this case? And you'll have to explain what caused you to think that this was relevant. And it's difficult to do that if you looked at a hundred cases without a road map.

Search histories and search maps can be used for many different purposes, as described later in this chapter. Many of the uses described in this chapter are based on the availability of a search history, possibly in the form of a search map. The search map also serves as a bridge between the physical acts of finding information and using information. The next section describes some of the characteristics of search maps, more detailed descriptions can be found in the next chapter: Findings: Search history data.

5.0.2.1 Characteristics of search maps

Process flow

One of the reasons for saving the search process is to be able to reconstruct the flow of actions, what was done, why it was done, and how the actions interrelate. In showing the sequence of user actions, displaying the types of relationships between actions is very important, as it greatly reduces the uncertainty about how the user got from one place to the next. This problem became obvious from collecting reactions from users to the Westlaw search history interface (Figure 5), which presents the documents, document set and queries visited by the searcher in a linear order, but omits information about the relationships between these documents. The lack of this information created a lot of confusion for users in the second phase of data collection, when they were trying to interpret the flow of actions from the linear display. Participants in the second phase of data collection often complained about not understanding the Westlaw search history screen, due to the lack of links between the search history elements. They could not find out from the search history how they got from one step to the next and thus lost the meaning of these relationships.


Research Trail - 03/26/2000 09:10PM for THE HAMBURGER, INC.

Name	Type	Information	Date
THE HAMBURGER, INC.	Open Trail		03/26/2000 09:10PM
MEAT BYPRODUCTS AND RESPONSIBIL...	Search	ALLFEDS	03/26/2000 09:15PM
Jefferson County, Ala. v. Acker	Document	119 S.Ct. 2069	03/26/2000 09:16PM
City of Chicago v. Morales	Document	119 S.Ct. 1849	03/26/2000 09:17PM
Board of Educ. of Oklahoma City Public S...	Document	111 S.Ct. 630	03/26/2000 09:17PM
Board of Educ. of Oklahoma City Public S...	KeyCite Citations	111 S.Ct. 630	03/26/2000 09:18PM
44 Case W. Res. L. Rev. 75	Find	44 Case W. Res. L...	03/26/2000 09:19PM
THE HAMBURGER, INC.	Close Trail		03/26/2000 09:26PM
THE HAMBURGER, INC.	Open Trail		03/26/2000 09:36PM
MEAT BYPRODUCTS AND RESPONSIBIL...	Search	ALLFEDS	03/26/2000 09:37PM
Jefferson County, Ala. v. Acker	Document	119 S.Ct. 2069	03/26/2000 09:38PM
334K11	Find	334K11	03/26/2000 09:38PM

Display Print Help Close

Figure 5. Westlaw search history screen from the web system. March, 2000.

Saving the process is not as easy as saving the results of the queries from a session. It requires saving most of the elements, or pointers to the elements, saving their attributes and relationships to other elements. This includes objects and actions.

When recreating context for an interrupted task, presenting the context as fully as possible may help explain why something was done. Seeing his or her own steps can be more meaningful to the searcher than a set of documents found, as it relates to episodic memory as opposed to abstract semantic memory. Seeing the flow of actions and relationships of steps to one another and their order can help users remember why they did something, or what happened in the search.

The importance of temporal organization lies in the fact that it represent the searcher's own path through the system and thus helps him understand the process, the relationship of events and the why-s of actions. It represents the map of the search. Overlaying the temporal roadmap representation on the information source structure emerged as a need in the participatory design sessions.

Personalization

Searchers' annotation on their search actions is another tool to record the thinking behind the actions; straight actions logs need to be augmented by user comments to make them more understandable and useful. This is especially important when searchers have to explain why they did something to help others understand their train of thought and action, as in the following collaborative task scenario. Coordination uses of search histories are discussed later in section 5.8.

I: And what do you do then?

S4: And then you would call up the documents. And the files.

I: Would you recreate your searches or …

S4: Oh no, I wouldn't recreate. just whatever is… , it's more just a history

> record. It could be that, if I have to justify a decision, which I haven't had to do but I could see it happen, something like, why did you give that opinion, and if it wasn't apparent from what I created, then I would have to do more search. I would have to try to recreate why I did it, why I said what I did.

User action metadata should be automatically recorded by the system, actions with content, linked to date and time and the name of the searcher. The duration of actions and the length of the session can be calculated from the automatically recorded data and presented when necessary. This information should be integrated into a search map representation of the searching and presented to the user.

5.0.3 Disadvantages of search histories

Search histories can be very useful in supporting information seeking, however their application in user interfaces should be weighed against their disadvantages for the context at hand. Information-seeking user interfaces are complex, adding search history information to then can lead to even higher cognitive loads for the user. Attracting to the user's attention to historical data can be harmful in distracting from the main task. When and what search history information to provide to the user depends on the domain, user tasks, user characteristics, system characteristics, and context.

In addition to distracting attention, presenting too much information on the search screen can be confusing to the user, especially to novice users. Search histories represent an additional dimension to information-seeking tasks, which are complex even without the historical dimension. Interpreting the various types of information presented poses a demanding task on users, showing another type of information can get overwhelming.

Search history information, in many cases, needs to be managed in addition to being viewed. These management tasks allow users to take full advantage of search histories, but they also add complexity to users' many information seeking tasks. While in complex, involved task environments managing search histories has its rewards, in other contexts thee management tasks would not pay off and unnecessarily burden the user.

In certain contexts, such as when searchers have unlimited access to databases, it is easier to reconstruct events than to manage and retrieve them from a search history. The effort of maintaining and retrieving from the search history is not justified for every task situation, and this should be considered for the system being designed and the user tasks in question. Some user tasks are simple enough that recreating them does not require complex problem solving. In this case, recreation is probably less demanding then retrieving the task from a search history. Users can spend considerable time constructing other steps, where retrieval can be easier than recreation. Thus, whether a search history supports or hinders users os a specific system should be decided based on user needs assessment data.

As discussed in Chapter 7, search system screen real estate is precious is terms of supporting searchers in information seeking. Whether search history information is displayed or not and what portion of the screen is devoted to it has to be weighed against the benefits it brings to users and the danger of distracting user attention from the information-seeking task and screen.

5.0.4 Legal and ethical concerns

In the legal profession it is especially important to consider legal and ethical issues in recording and storing search history information. States have rules on how long law firms are required to store records from their cases. These need to be considered when search history data storage systems are designed. Law firms and corporation also have internal policies that determine how long and in what format the records need to be stored. These rules and regulations need to be followed when creating search history systems.

> I: Ok, the last set of questions is about, if you can add anything about using search histories in searching.
>
> S4: Well, the only thing I could add is that generally within corporations we have guidelines on keeping documents. I have files that are not relevant, that I feel will not come up again, then I will destroy. Automatically. But if it could be relevant in the future, then I will file it in the corporate archives and then it's kept for like five years. And then destroyed.

Another contextual factor influencing search history and results archiving is the threat of malpractice suits. It is the professional responsibility of lawyers to carry out exhaustive information searches in order to best represent their clients. If they miss an important piece of information, they can be prosecuted for malpractice, as described in the following example:

> S2: lawyers, we often think everything is important, every little paper must be kept forever.
>
> I: Why is that?
>
> S2: because at any time you can be sued for malpractice.
>
> I: that's an important point.
>
> S2: You need to be able to always defend yourself against your clients, which is sad.

The automatic recording of search histories raises privacy concerns that need to be managed but provides an abundant source of information for reuse. Often the user cannot anticipate a need for search history information at the time of recording and may need data that was not recorded. In addition, the automatic analyses mentioned above provide rich reporting tools.

5.1 Memory support

The most elemental reason for recording search histories is to support human memory by externally recording information from user actions and system responses. Information from memory is used in many different ways, just as the external representation of memory is used in many different ways. This section refers to the theories exploring the workings of human memory described in the Literature Review in section 2.1. Cognitive Background; these are often referred to in the paragraphs below.

5.1.1 Working memory support

Searchers often needed to compare similar items: run very similar queries, compare result documents, state whether they have searched various sources, etc. In these comparisons, searchers must attend to two items of one category and attend to them long enough to be able to make comparisons. Searchers jotted down notes, printed result lists and documents to help them with this task. One example of this activity is comparing queries in order to keep track of progress and avoid repetition.

> P1: I can't remember how I, I know I put separation within so many of church and within so many of state, and I thought I put “and textbooks” “and public board” “and public schools”, but I can't remember. I guess if I just kept going back forever?

Keeping all items in a history record allows the searcher to display these next to each other and makes comparisons easier (working memory aid). The system can also show the differences between similar items to help users. Showing repetition of items is an even better solution, as the system can keep track of queries, documents, etc. and show repetitive display and use of these automatically.

Chunking helps working memory, which holds only seven ± two units of information. By providing a clipboard to users to organize and group items from their search history, we can help them with chunking by physically representing the chunks. Spink et al., in their study on reference librarians searching notes' found that searchers use many graphical elements in their notes. Some of these graphical elements serve to represent relationships and groupings. The clipboard tool can support both of these functions. One of the participants described clustering of results found as the first step to grouping and structuring ideas for the brief she was writing:

> I: The clusters of documents or groups of documents were also ranked high. When you find cases, would you want to group them into your own categories and keep them like that?
>
> P2: Yes, I think so, particularly if I were writing a brief, I would probably, I might want to break it down into how I was going to divide up the brief and which case is going into what subject area, so that would be helpful.

Creating groupings also provides deeper levels of processing, which support learning, as discussed below.

5.1.2 Episodic and semantic memory

Episodic memory comprises particular experiences dated and timed and placed in space; semantic memory is more abstract, it is a human thesaurus of concepts, terms and their relationships, it is abstract knowledge. (See Literature Review section 2.2.4 Episodic and semantic memory) The theoretical distinction between episodic and semantic memory has very interesting manifestations in the data collected. Participants differentiated a record of what has been done (specific experiences with the system and the data), versus what they have learned (abstract knowledge of concepts and relationships).

Search maps described by searchers seem to represent episodic memory, the searcher's exact path through the system, representing all the events and actions. This type of information is useful when the searcher has to recreate an action sequence in order to share it with someone else or to remind her of what exactly has been done and why she made certain decisions. This is necessary when she has to account for her actions and decisions. Planning ahead, rather than staying on track.

Semantic memory is defined as abstract knowledge of concepts. While the searcher reads and interprets documents, she learns about legal issues and builds this new information into old representations. Various knowledge representation theories exists; for the purposes of this dissertation, mental model theory is used to describe long-term memory and knowledge structures.

In practice, however, mental models and search maps cannot be clearly distinguished. Search maps (called road maps by many participants) often help users interpret information and thus provide a bridge to forming mental models, as illustrated in the following quote:

> S2: Road mapping.
>
> I: Can you elaborate on the road mapping?
>
> S2: You know I don't really know, except for an issue comes before you, how you think about the issue, how you turn it over and manipulate it, can influence your search results. You need to make sure that you are manipulating the issue in the right way. A lot of times you can see, as a young associate or junior associate, you go to the partner and they'll say why did you pull this case? And you'll have to explain what caused you to think that this was relevant. And it's difficult to do that if you looked at a hundred cases without a road map.

5.1.3 Level of processing

Note-taking represents a deeper level of processing than simply reading a document, as it involves possible reformulation of the text. According to the level of processing theory of memory, deeper levels of processing will help remember and retain memories for a longer time period. Providing tools to process results, take notes and annotate can help users learn new information more effectively. One of the participants described using notes to help learning and remembering:

> P1: Typed annotations, I forget things all the time.
>
> I: But you wouldn't keep it, you would keep it for a session but not…
>
> P1: Yes, I wouldn't keep it, just the way I work is, by the end of the session, I'll cement it into my thinking hopefully. But I guess maybe you are right that if you have one reason or another you might want to keep those annotations for more than a session.

5.1.4 Types of forgetting and the potential of search histories

5.1.4.1 Aid with attention and working memory

Search history can help with loss of information due to failure to attend. Transfer from sensory memory stores to short-term or working memory is achieved through attention. In a rapidly changing user interface environment/context it is difficult for searchers to pay attention to and remember specifics of searches. Increased availability of information involved in searching (queries, results, documents, groups of documents, steps, etc.) gives users a second chance for attention, thus supporting transfer of information to working memory. In busy search screens there are too many things going on at the same time, searchers only focus on part of the screen. They may later realize that they should have looked at some other part and need an easy way to go back.

> S6: Oh, so there were things that I did not see possibly? Is that correct?
>
> I: Probably, I can't remember exactly.
>
> S6: Let me just take, before I get to, this is that ALR article, is it? I'm just curious to look at what the five pieces were and how they broke it up to see if it is of help. What would be helpful here, is maybe a way that they could show me, just like an outline of the file parts. And maybe they did and I thought I did not see what each one of the screens was going to do for me. Because here I may know that I want to go to screen four. Instead of having to go and look through the whole thing, if there was maybe a master plan that maybe helpful.

Presenting queries on search results pages, displaying all queries on the formulation page and other interface design tools can help users in going back to previously viewed screens, as well as the presence of a search history trail on the screen.

Focalization helps humans select a set of inputs to pay attention to from a myriad of events and information. It helps us cope with information overload. Decisions about the importance of information made at the time of the event are influenced by the current knowledge state of the user. Reevaluation of attention decisions may be needed later on, as the searcher knowledge state changes, (delayed relevance). In this case the users must go back to a previous state and focus on different details. Search history allows the searcher to go back to previous system states and focus on a piece of information previously not attended to.

Keeping this information on the screen and in the sphere of the user's attention can help searchers make sense of the search by being well informed about actions and results. It can also be helpful to present notes and goals on the same screen, as discussed below. Presenting information in its context can help users with interpretation. At the same time it is dangerous to overburden the user with many distracting factors, it can be especially complex for novice end users. The choice of displaying the history trail or not should be left up to the user.

Participants in the study often switched back and forth between their notes and the screen. This observation led to the conclusion that while searching, searchers needs to look back at their planning notes to remind themselves of searching criteria. Presenting this information on the screen seems to be important, although it takes up screen real estate. Decisions on what part of the input to pay attention to are often based on how it relates to previous knowledge. In searching environments, an initial analysis is carried out in order to focus attention on information that is potentially relevant to an information problem. Representing the original information problem, the thoughts and notes of the user while he is reading and interpreting results may help him focus his attention to relevant information. Presenting search history in a separate window will likely to distract some attention from the main screen, but also helps the searcher keep in mind what to look for.

Based on attenuation theory, we can assume that presenting selection criteria on the screen in the search history makes it easier for the searcher to select information from the main channel. Although the SH is only the secondary channel in processing information, important words are picked up from it and it is compared to the primary channel, thus helping the searcher tie in the new information to his knowledge, tasks and goals and information found or created in previous steps. An example of referring back to user notes or search histories is described in the following section, where the searcher uses a different application with a history tool:

> I: When you use this application, do you use that trail list?
>
> S5: Yes, I do. Because I forget whether I looked at something or not. And then it'll show the four words that I have looked at already in any given session. So as the day goes on, as I keep searching things I can look back and see. And then I could go back to, this is the second search I did, and I went back to that by going to Trail. Let's say I wanted to look at 'bath' again, I'd go back to it. If I want to go forward to something, I'd go forward to it.

In addition to user notes, search histories contain previous steps of the searcher and system output, result lists and document. Showing new information in parallel with previous information may help humans focus their attention to that which fits best their situation through highlighting the relationship of the new information to the old.

5.1.4.2 Aid with encoding into long-term memory

Search history can help overcome ineffective coding and failure to encode. Encoding is the process of moving information from working memory to long-term memory and storing it in the knowledge structure of the user. This is usually achieved through working with the information one way or another, as discussed in the level of processing theory. Manipulating queries and other search and browsing steps, organizing, and annotating results can help users learn about the process, the system and the subject area by allowing users to more deeply process and analyze information. Taking notes involves reformulating information, it is one way to manipulate and reformulate information that helps users not only to record this for future reference, but also to better commit these to memory. Forming relationships between new information and information already in the searchers memory supports learning. As described above, one of the participants mentioned that she would only take notes with the explicit goal of helping her learn new findings and she would destroy the notes after the search sessions.

Other manipulation tools also support the searcher in deeper processing of the information. A clipboard tool, result-processing tools, and history representation are examples.

Only part of the information is encoded, other parts are stored in external stores. Often there is not enough time during searching to devote enough processing to learn and remember the information. It may also be impractical to try to remember all the information while searching, as this may interfere with the process of searching itself. Participants often committed to memory only the information necessary to continue the search (although even this may exceed human ability) and stored other information for later processing in external stores. Examples of these are search histories and results gathering and processing tools. One of the participating attorneys described this in the following way:

> I: Why are you copying it? Do you usually do this? Do you usually copy the text.
>
> S6: So I can read it.
>
> I: So that you can read it?
>
> S6: Yes, I can't … It costs too much to stay online and to do research. I take usually the best of what I find, I either download it to disk or print and then go back and read hard copies anyway. In the past you needed hard copies a lot more than you do now. They are becoming less and less necessary, but the search time, I like to keep going with the search. I don't like to get bogged down by reading,

Saving cases for later reading can be motivated by two different factors, one of them is the cost of online access time. When the searcher is charged for the database use by the time he spends online, it makes financial sense to download the cases and read them offline. Another factor motivates searchers to avoid interference with the main task of searching by reading cases in detail.

5.1.4.3 Aid with retrieval from long-term memory

Search history can help overcome failure to retrieve due to interference and cue-dependent forgetting. Even if the information was encoded properly, retrieval may be problematic. Retrieval problems include interference and retrieval cue failure. If a search history exists, this information can be more easily retraced. This is also true in the case of interference. Interference suggests the blocking of information by new information learned (or old information).

Another problem that can cause forgetting is that the right cue is not presented to retrieve the information. Search histories can help this problem by presenting multiple context elements from the information, if not the information sought itself. Encoded information is retrieved by cues that were remembered at the time of encoding it; presenting the context of events in a search history can be helpful, as the contextual cues are tied together by time and location. One example of this is remembering various cases. It is easier for people to relate to the human story in a case, the facts of a case, as opposed to the findings or the specifics of legal issues. The case about the teacher forcing the kids to sing religious songs in his class is more memorable than the decision on separation of church and state and the issue of establishment clause. Presenting facts from a case may jog the searcher's memory better than presenting the citation or the decision and enable him to retrieve other specifics from memory. Several study participants described this, as the following example shows:

> P1: I think a lot of times you remember cases not by actually what the holdings are, but the facts of the case. Oh, that was the case where the teacher forced the kids to sing religious songs. And that would usually come out in an abstract, just a little a bit about the facts that are distinct. And that's how I … and then the fact that I remembered that it's a religious song case, then you remember why it's important to you, and then maybe, probably you'd remember what the holding is. But if somebody just says a case where they held that you can't …

It is also important to allow searchers to name units of their search history, create markings and annotations attached to the search history as they may pick out additional cues that will remind them of the information. Highly relevant retrieval cues are the most important for cued recall.

Recognition from a search history is less demanding then trying to recall the information from memory. Presenting highly relevant information can help searchers recall other information and thus the search history can reduce cognitive load.

5.2 Search system use and problems of navigation

5.2.1 Search system use

Search system use involves tasks associated with the use of the search system as a computer application, in this case the Westlaw interface. Tasks include starting and closing applications, finding functions, using functions, transferring information between applications, and so on. System use and within that navigation is an important use of search histories. Activity histories in general are useful in application use; one example of this is the undo/redo function. The data collected show that hypertext information seeking and use can also benefit greatly from the application of activity histories.

An important difficulty is caused by the dual metaphors participants refer to when using web applications. They often think about functions in web applications in terms of going back and forth in a linear pattern, applying the hypertext navigation metaphor and ignoring their previous familiarity of PC based applications and interface tools such as menus. A searcher would try to "go back" to the search screen to create a new query, often having to go through several screens. He would not look for the "Search" menu item where he could access the same screen within one click. This discourages users from using menus and thinking of finding functions in terms of "going back to the search screen" as opposed to finding the function. This may also be influenced by the fact that users want to get back to a certain function with certain content within that function (e.g. a previous query in the search screen).

There are many areas of application use where activity histories can be helpful. Three of them emerged from the data as especially important for hypertext search systems: navigation, finding functions, and undo/redo. Repeating previous actions from a search history can remind searchers of the syntax of those actions. It helps with system learning and serves as an example to help the searcher.

5.2.2 Navigation and disorientation

To be able to navigate the system, users must know the answers to the following questions at any given time. These questions were collected and grouped from participants' own questions while searching for information:

- Where am I? What is this?
- Why am I here?
- How did I get here?
- Where did I come from?
- Where can I go from here?
- How do I get to X from here?
- How do I get back to where I was before?

- How do I get back from a tangent I followed?

If she knows the answers to these questions, the user should be able to get around the system and find functions necessary to solve her problems. Navigation and disorientation are especially problematic in hypertext systems, where it is very easy for the user to get lost. The legal literature has many citations among its documents; this is often implemented through hypertext links, as in the Westlaw system. This led to problems of disorientation and navigation. Search histories can show the navigation path through hypertext document sets and help users find their way.

One of the most frequent uses of search history in navigation is going back to a previous place. The searcher must know where she came from and how to get back to that. By providing a clickable history, searchers do not have to remember this information but can simply recognize it from a list.

Navigating is helped by creating landmarks. The user should be able to create landmarks that help them remember and find their way back. Especially in cases when users know in advance that they will return, as when following a tangent. For a more detailed description of this problem see the section on tangents.

Disorientation does not only hinder navigation, but knowing where the user came from and how he got there may help him figure out why he was there and what he was doing.

> S2: I don’t know where I came from there from the footnote? Where did I come from? Now I have to go back and figure out how I even got to that Law Review article in the first place and why was I there. I forgot. … Right, now once I figured out where I was I know what I’m looking for. I’m looking for annotated cases that are interpreting a certain part of this statute.

Hypertext disorientation is a widely researched problem and many suggestions have been made building on a track of user actions. Showing users their paths in the systems superimposed on a representation of the system structure may help with this problem, as it can help users understand why they went from one place to the next. Implementation of this combined display can be complex.

In order to help users take advantage of search histories in quick navigation decisions, the history record should be available in an easily accessible format. This is especially helpful when following tangents off the main task avenue, as described in the section on tangents, section 5.4.5.

5.3 Find information/Search

Search history information can be helpful in many information-seeking tasks. The uses of search histories described in some of the other sections of the chapter ultimately also serve to support searchers in finding and using information. The current section focuses on direct searching and browsing techniques and actions.

5.3.1 Information seeking as problem solving

As discussed in the literature review, information seeking is a problem-solving task. The stages of problem solving identified in the literature review were the following (Hayes, 1989):

(1) Identifying the problem

(2) Representation of the problem

(3) Planning the solution

(4) Execute the plan

(5) Evaluate the plan

(6) Evaluate the solution

The system can provide support for these various stages of the process and thus facilitate arriving at a solution. Identifying the problem or information need is challenging, as discussed in the next section. The information need can change as the search progresses, or the searchers attention can be distracted from the goal. Creating tools to record and display the goal and organize results by subgoals can help searchers better understand and define their needs.

The ability to explicitly represent the problem and avenues of solution through interface tools can help searchers identify the right solution. In case of information-seeking problems, keeping track of various pieces of information, the ability to compare and manage these can help searchers. Searcher and design participants in the study both described planning actions in their information seeking. Planning is taught by instructors of legal research, as well. Tools to support planning are needed, as described later in this chapter. Encouraging evaluation of the steps and the results can prevent the use of poor quality information. Tools that represent the steps of a search sessions will help with this evaluation. Evaluation is also discussed below.

5.3.2 Planning/information needs clarification

Reference librarians are trained in helping patrons clarify what kind of information they need through a series of questions and discussions in a reference interview. When the searcher is alone with his computer, tools to help him clarify what he needs are important. In legal research some of the dimensions are given (e.g. document type, kind of information) but others, especially topical ones can be supported through history-based tools. End-user legal searchers often go to a database to explore a topic and find out what the definitions are, how the area is structured, and what issues they need to search on. Capturing this exploratory searching, along with pre-search notes, can help searchers learn about an unknown area and better define their topics. User notes can help by providing a continuity of topics, creating a space to save sources, keywords to search. In the following two examples the searcher describes how she uses her notes in keeping track of topics, and keywords when refining these through the searches. In this case the needs clarification is strongly linked to the search plan, as the searcher looks for the information right away:

> Halvorson: Do you file your notes so that, if the case file comes back, you could readily find your notes for it?
>
> Best: […]When I first get the project and map out my initial strategy for where I want to look, I put the strategy in that folder. I add to this as I refine my research strategy. I keep a rough outline of the issues I am researching, and flesh it out as I go with sub-issues. As I proceed, I include notes on which cases and statutes I've updated, and notes for every source I've looked at. I put material specific to a discrete issue into the sub-folder for that issue, and cross-reference material that relates to more than one issue.
>
> Halvorson: How much planning do you do before logging on to QuickLaw?
>
> Best: I do a fair amount. I always make a list of keywords, synonyms and alternate words, and think about which words should be truncated. I think about what databases to search and make a note of them. Then, as I do the search, I refine it and check off what I've searched. Depending on how complex the search is and how many alternate terms it includes, my search plan is more like a diagram, with several columns of alternate terms separated by the appropriate connectors.

The system can help with clarifying the information need with assistance in query formulation throughout the search.

Keywords are often part of pre-search notes. These are very characteristic of the user and his areas of interest, and have a role in guiding the search, as in the following example:

> S6: Keywords, very important, all along, because those are the guiding, organizing principle.
>
> I: Notes?
>
> S6: You saw my notes.
>
> I: Yes, it's usually not continuous text, it's keywords, reminders.
>
> S6: yes, I usually go into searches with keywords. That's the one thing I try to think of because searches are all keyword, or mostly keyword-organized. Phrases, or concepts, sometimes citations, but it's really the keywords or the subject matters will narrow down to get the closest thing of what I'm looking for.

These keywords will transfer to queries, define queries, when searching starts. They can be used to create task representations and monitor progress against them. Systems should support this transfer from notes to task representations and queries, and finally search histories and completed task lists.

Keywords in queries represent the content of searches and strategies tried; they are central to the query formulations. In the next example the searcher is collecting keywords to try in future queries:

> Halvorson: How much planning do you do before logging on to QuickLaw?
>
> Best: I do a fair amount. I always make a list of keywords, synonyms and alternate words, and think about which words should be truncated. I think about what databases to search and make a note of them. Then, as I do the search, I refine it and check off what I've searched. Depending on how complex the search is and how many alternate terms it includes, my search plan is more like a diagram, with several columns of alternate terms separated by the appropriate connectors.

> S6: Malpractice cases … [reading]
>
> S6: That's a little bit more on target. Let's go take a quick little look at this. That's a word that we could save now, that I'm going to write down in order that I may want to use as a search term. That I saw in one of the cases that they used. It's not too new, it's just a different way of saying something that's probably more safe for doing a more formal search.

5.3.3 Development of search strategy/Query development/Execute searches

The development of a search strategy is a complex task. In initial planning users may jot down topics to search for, however, during searching they need to determine many more variables such as source, query terms, query syntax, results format and so on in a lot more detail. Selecting and combining these variables to form search strategies and queries is a challenging task. In the non-virtual environment of human actions, we often see, hear or experience things that we will build on in the future. Most human actions are based on learning these actions at one point in our lives and by learning more and more about our environment and its affordances, we are more successful in operating in it. This is true for the digital environment as well; previously encountered information has a role in future actions. The systems can record this information and thus give a better account than the user's own memory. In searching users may want to reuse previously constructed queries in the same form or with modifications. They may want to use keywords or sections found in results, or use a set of bookmarked sources.

End users and expert searchers have different levels of knowledge and experience in searching and use different tactics to accomplish their goal. However, both groups can build on search histories as a support tool. Plans and reminders can be recorded in search histories, and the progress of the search can be compared to the plan. These two functions are further discussed in the section on task management, section 5.4. Searchers constantly evaluate and modify search strategy; search histories help searchers review the actions and results and modify their plans accordingly. Comparing sources, techniques, and documents side by side is easier when access to previous steps is easy and their results were recorded. Comparisons are discussed in section 5.5.8.

The following sections will discuss one specific use of search histories in query development: the reuse of previous information in creating new queries.

5.3.3.1 Reusing search actions and results

Search histories make searchers aware of previous actions and object in the search and create easy access to these. Reusing previously recorded information saves time for searchers and allows them to create more efficient searches. Attorneys and search professionals have different emphases in reusing information. The participants in the study included both attorneys and law librarians. While librarians would most often reuse queries, sources and strategies, attorneys tend to reuse the content of results or their own documents. This is easily explained by the difference in their tasks: librarians focus on finding information and need tools to accomplish that, attorneys work with legal issues and need solutions those – which is described in documents. In some contexts it is useful to facilitate attorneys' reuse of queries in searching. The following sections describe reuse of historical information, either in the same form or with changes.

History-based tools can further help with clarifying what it is the user needs and where to find it by analyzing what the user has been doing and make suggestions based on the records: for example, if the user looked at case law but has not checked law reviews, the system could suggested relevant law reviews. Or the system can suggest terms based on a thesaurus. This need is described in the following quote:

> I: How do you know when you're done?
>
> S2: You go until you keep seeing the same cases over and over again. That's a talent. There are a lot of senior attorneys who just know when they're done. And associates can spend hours and hours, and days and days and just never feel done, because you worry that you missed something.
>
> I: So you just go on until you keep seeing the same cases?
>
> S2: Or I think in Westlaw there are prompts to tell you other things that you may want to look for. That's why the keycites [headnotes? keynumber?] can be very helpful. You know the name's different but the keycites may be the same, that's if the person at Westlaw knows what he is doing. Once you start seeing the same cases, you figure you exhausted it.

5.3.3.1.1 Reuse same query

Participants in the study often reran the same query for any of the following reasons:

(1) The searcher wanted to see a set of documents he had worked with before

(2) The same topic needed to be searched in a different database

(3) The search results needed to be updated

(4) The searcher wants to share the query with others

Rerunning a query can be helpful in recreating the context for an interrupted task, where it is important how the thinking of the searcher progressed and what the natural next steps are.

Rerunning the same query occurred in the study when participants wanted to look for the same information in multiple databases or sources. The Westlaw system used in this research organizes information into clusters: over 10,000 databases. In order to find everything needed, attorneys often must run the same query in several of these databases and present the results separately. This can facilitate learning about the coverage and structure of the different databases as well. The following quote illustrates running the same query in a different database:

> I: Write your query down? Do you write anything else down?
>
> S5: Actually usually what I do is I write down the query that was most successful, I might not write down all of them, but I would write down the one that was most successful, the one that got me to the information that I needed, so if I have to do the same query in a different database, I got it right there… I would make sure that I write down, maybe at the top of the paper here [the printout of the result list] what I had searched in order to retrieve this. Exactly as I entered it.

In litigation, queries need to be periodically updated to see if the law is still good law and what new developments occurred in the area of a legal issue. This is crucial for successful litigation and is an ethical responsibility of the attorney. This serves similar purpose to citation checking in section 5.3.4.3. Updating searches can happen through rerunning the same query, though in some cases terminology changes very fast and terms may need to be updated in the query. Rerunning queries for the purpose of updating the results is exemplified in the following quote:

> S2: If it was more than a week, you still have to go into a search and make sure it was current. […]
>
> I: And then when you go back and update how do you do it? You take your research file and you look at what you searched on? You keep the search history…
>
> S2: Yes, because your search history is in there.
>
> I: And then you repeat the exact same search?
>
> S2: Or it’s the terms have changed, you know, yeah. But that’ll depend on what it is, there are so many different ways to look at it, some things you can go to the newspaper databases, clip services and see, because the danger with using Westlaw and Lexis is when’s a couch a sofa, when’s a love seat. When same-sex people are gays and lesbians. You know you miss a term, you missed the case. I call those sofa, you call it couch, a divan, are you really going to

> know those all?

It may also be a way to reach previous results through rerunning a query. It is often simpler to get results back this way, as opposed to save them locally and retrieve them later, or hitting the 'Back' button of the web browser multiple times. This distinction can be hidden from the user through recalling results from a history (thus having the opportunity to preserve annotations and other personalization features) and updating it from the database.

Rerunning queries and searches was used as a tool for collaboration, to share the results and the process with others. One participant described replaying or displaying a set of steps to help explain why something was done as a method to train others.

> I: So here you can go back to Lexis, when you go back, would you like something there that's personalized?
>
> S3: […] Or if you're working in coordination with other attorneys and you want to show them or just tell them the searches you have already done, or you want to show a search that this is something you can do or this is a search I have done to help others learn, then yeah, I can see it would be helpful.

In a different scenario, experts publish queries for less experienced searchers to use. A law librarian describes a service that she finds useful for novice searchers:

> Halvorson: Have you ever looked at published electronic queries designed to, say, update an ALR annotation?
>
> Botluk: […] For people who are not comfortable with online searching, those published queries are definitely good tools.

This type of search history represents a collective memory of good techniques used in the domain of the organization's work. These published queries are similar to FAQ (Frequently Asked Questions) services, where the organization anticipates questions often posed by users and publishes these with the answers provided. The queries shared are usually ones that are typical tasks in the work of the organization.

5.3.3.1.2 Modify query/Reformulate query

Most attorneys are not experts in searching, like law librarians are. They use many end-user search strategies, for example, many trials while formulating a sufficient query. The process of reformulations are part of the information-seeking strategies, as described by Bates' berrypicking (1989) model.

Exploratory techniques are representative of end users; participants in the study often reformulated queries and tried many versions before reaching satisfactory results. A history of previous actions can support reformulation, planning and evaluation of the steps. Participants used an empirical process to learn about the database, searching techniques and topics. They would send out the first query as a feeler to find out if they are using the right terms, if the database has content on the topic of interest and if so, how much and also to get ideas on how to proceed. If a topic is new to the searcher, the first query is often feeling around in the dark; if the topic is familiar to the searcher, the first query is based on previous searching and the user has some expectations of the results. Organizing the search results also helps the searcher with exploring the topic; see Section 5.5 on search results management.

The examination of the results set returned and optional interpretation of the results can show users how to modify the query in order to achieve more useful results. The set may be too big or too small, it may reveal a new use of terminology that requires changes to the query. Finding new terms to use is a typical example of reformulating a query based on the results returned to the original query; the following comments are examples of this type of query reformulation:

> P1: Ok, I first searched on natural language search and found a lot of cases that weren't particularly relevant but by looking at those cases I was able to find different terms that I should have used in my search to begin with. So I amended my search to reflect those terms.
>
> I: You start reading, you start learning more about it? And then what do you do next?
>
> S5: Then what I would do is look for similar terms, the law might be known by a different name as well. Sometimes laws are known by their formal name, as well as a more popular name that comes up later.
>
> I: So you would find something and you would search on that as well?
>
> S5: Right, I would search. If this law is known by a different name or by a different phrase, I might search that other name or phrase. And possibly come up with more information that way.

Reformulating a query is a form of reuse of a previous query with modifications; good edit capabilities are needed to support this task. Also the searcher needs a good record of multiple reformulations in order to keep track of what has been tried and what is still planned.

Searchers read documents while querying the system in order to find new search terms to use, to identify important authors in an area or identify new facets of the problem. In the following example the searcher reformulates a query based on new terms found:

> P2: I think I would look more, I would go back to my first search with the textbooks and the separation of church and state, establishment clause, that was one thing I found, separation of church and state wasn't a good way to search, establishment clause was much better. And look for law review articles, very recent law reviews on establishment clause and maybe public school to see if there is anything like that. I would keep looking at these school, public school cases, there were 140 cases and I've always been somebody that trudges through things, so I would just keep trudging through things.

There are several typical tasks when searchers need to rerun previous queries without modifying them, but other times the system can help exclude erroneous repetition of queries, as the one shown in the next example. Here the searcher explains taking notes of queries in order to avoid repeating them:

> I: What would you write down?
>
> S5: I would write down exactly what I had searched and in what database I had searched, so that I didn't repeat myself, so that I didn't do the same thing twice.

As the many reformulations may result in a high number of similar queries, it is hard for searchers to distinguish and remember what queries they have run already, and this often leads to repeating queries or confusion. Highly similar queries form natural clusters in users' memories; the search system can compare queries and cluster them automatically. Chunking helps humans' working memory. Providing a clipboard for organizing and grouping items from their search history helps searchers with chunking by physically representing the chunks.

A special case of query modification is combining previous queries, which is often a function in search systems. This usually involves a list of previous queries with the option of setting up AND and OR relationships between them and rerunning them in the combined form. A need for the NOT logical operator in combining queries is described in the following remarks, where the participant would have liked to exclude the results returned to a previous query:

> S1: You know what I should do? Is there a way when I can do ALLFEDS but not 8th circuit? That would be nice, because I just looked at all that.
>
> I: I don't know.
>
> S1: It would only take out 12 cases but…
>
> I: So you wouldn't want to see the cases you've seen before?
>
> S1: Yes, if there's a way that I can exclude what I just did … well, it's too bad. But it's only 12 cases, so it's not that bad. It would be nice though.

Combining previous queries is especially important for searchers using building block search strategies, where searchers can start with a broad search using all the terms identified as pertinent to a topic and then narrowing it by dropping terms if the returned result set is too large. (Armstrong & Large, 1988) Alternatively, the searcher can search separately on various terms and afterwards combine the queries.

Searching within a result set can be achieved through a modification of the query, though most users do not interpret it that way and instead want to manipulate a new target document set. Thus this topic was raised several times in the study by participants, and is discussed in the section on the processing of results.

5.3.4 Search techniques

The type of query can influence the need for recording search histories. For simple fact finding or known item searches the result may be the only important information to record; saving the process, other than possibly the source, would not add to the information seeking. On the other hand, complex searching tasks can take advantage of the process. The information needs in this study were all complex, served by topically-driven searches and exploration.

Search techniques used also influence the use of search histories, and at the same time are shaped by them. The tools available to searchers affect the way they look for information. There are several strategies taught in law schools and in library schools to be used in legal research. These influence searchers' practices and the way they use search histories. Strategies taught often allude to names of sources, range of sources to check, order of sources, types of actions, etc. The most important techniques observed that built on search histories were exploratory strategies, citation chaining, and high recall searches. Exploration of the search space was described above, in section 5.3.3.1.2.

5.3.4.1 Citation chaining

The highly intercited nature of legal literature encourages a lot of citation following from cases. When a good case, article or a relevant annotated code is found, searchers would often use that to follow all the links from it to look for further information. Citation chaining is an often-used tactic in legal research, as case law, annotated code and law reviews are all highly interconnected through citations. These connections are represented by hypertext links in the Westlaw legal databases, where following citations is even easier. So that the searcher does not get lost, the original document should be saved and marked in an easily-accessible place, a hub where searchers can return to it to follow new links. The original case is used as a plan or a checklist of actions and should display the searcher's progress. When following citation links, navigation and disorientation often caused problems for searchers. Providing search history displays to present their path back to the original hub document can help with this problem.

5.3.4.2 Citation checking

Citation checking (not the same as citation chaining) is defined as the updating of cases over time; it is an important step in legal research. Litigation in a case can stretch over several months or years. Cases found in searching at the beginning of the process and used in arguing can be overturned during this period of time, in which case the law defined by the case changes. Attorneys have to stay up-to-date on the changes in the outcome of cases used. Often months or even years can pass between the original research and the actual use of the information in trial. In this case, all the citations used in litigation need to be checked for updates periodically, as the trial requires. The study participant in the example below tells a story about the importance of citation checking:

> S1: One thing I didn't do before I printed it is Shepardize it which I typically do, yeah, I always, because you never want to get into a situation when you print out a case, then you find later on that it was overruled, I had that happen once when I wrote an entire brief and built it around one case and I had Shepardized it, or at least I thought I had done it, and about six hours before I filed the brief, the paralegal who was doing the final cite check on this found that this case has been overruled. Actually it has been vacated on other ground, so I could still use for the point I wanted to sue for, but I thought I was going to get a heart attack. And I didn't do that here, which I normally do, but if it's something, for example, if I'm doing research and I come across a case which I think at some point may be one that I'm interested , I may print it out or I may just create a file for you know miscellaneous research or something like that, I have not started to maintain electronic files of those kinds of cases, that's actually something I'd like to do…

Citation checking should be easy to do or even automatic, based on previous search and information use history. The user should be able to select cases for future citation checking and receive notification if the status of the cases changed.

5.3.4.3 Observing repeated occurrences of cases

Another strategy that was observed is looking for cases that the searcher has seen before. This is an example of direct use of the search history, where the user is interested in cases that are part of the history. Cases that have been seen before may be interesting for a variety of reasons; one of these is that if a case is returned to several of the user's queries, it is likely more relevant than other cases. Another reason can be that if a case has been seen on several occasions, it is likely an often cited case. Showing repetition among results and queries is important for use. In the following example the searcher describes recognizing a case that he has seen before which contributes importance to that case:

> I: When you say you're looking for cases that you recognize, why is that?
>
> S3: I guess I'd, if there is a case that I have seen before, it's more than likely used a lot I'd be more likely to use that, if it's been followed quite a bit.

In the following example repetitive emergence of the case shows that it was relevant to multiple topics of interest to the searcher:

> S6: I'm going to take this case because it seems to have appeared on the radar screen in a couple of areas I was interested in. I recognize it from… I saw it in a different section of the Keycites before and it seems to be, well I'm not going to read the whole thing right now, it seems to be a case that may be a …

A case that occurs again can also be helpful in navigation: the searcher can easily jump to the results of the previous search that retrieved the same case. Filtering out repeating cases is also important when collecting results:

> S2: Well, yes, if there was a way to put everything you did into the file, and then print it when you're done, instead to download to print. Download to print just takes time away from your search process. And then to be able to go back and see, oh, something that comes up and says you downloaded this case three times. You know do you want to print all three? No, you just want the one. Yes, it's going to tell you what you downloaded to print.

5.3.4.4 Update results of a previous search

Litigation and other legal tasks often stretch out over a long time period. Searches executed and documents found at the beginning of the process need to be updated periodically in order to verify their accuracy. Searches need to be updated in order to find any new developments in the area, while court opinions about cases referenced can be published that reaffirm or overturn previous decisions, thus endangering their use for the task of the searcher. Rerunning saved queries and checking documents are frequent tasks in the legal information domain. Saving documents and searches and marking them for update can help with this laborious task. The searcher can be prompted to run the updates, or the system can automatically detect and report changes to the user.

Although sometimes searches and documents are saved to be updated, in reality they are often not. Tools for deleting, reorganizing these records should be part of the interface. Reminders about updating can be created based on different time periods in order to help the user manage the updating process.

Updating queries and searches has been described in section 5.3.3.1.1. Updating of the search results is most often carried out by someone other than the original searcher. The attorney may do the searching itself and then delegate updating to the legal assistant. Sharing the search history for updating purposes and then reporting the results back to the supervisor is an important part of this process.

In some cases changes may be needed to the search history when updating it. Terms and vocabulary may change; the focus of the information problem may shift, and the alterations in the queries reflecting these should be tracked. Keeping track of earlier versions can be important for various legal tasks, as described:

> S2: If it was more than a week, you still have to go into a search and make sure it was current. I'll tell you always, I mean your research is always flowing? But you never throw anything out, because legislative history and positions through history are very important.

The updated results should be incorporated into the records of the previous results. In the following example the participant describes updating his topic-oriented files:

> I: And if you get a new issue and you go back to your old files, do you update

> that? Do you go back to search?
>
> S2: If there is something that comes up, big change of rule, used to be able to do this can't do this anymore, I go back to the old file, put a sheet of paper in there and say, in October 1996 the National Securities Market Improvement Act was adopted. And it affected suitability, state regulation of insurance products and leave the file as it was and then you start a new file. And then you also file separates, because then you need to go back and see. And every time an issue comes up that was affected by that you have to go back to the file to see how it was changed.

The requirement to update emphasizes the importance of recording dates and times for search history and results records. In order to compare dates and times, the original should be recorded.

5.3.5 Relevance judgments

The first step of results processing is judging relevance. This involves a decision about the usefulness of the information found to the user's problem. These relevance judgments are based on reading and interpreting the results and relating them to the searcher's own knowledge and problem at hand. The results that were deemed relevant enough are saved for future analysis or use. Search histories can help users making relevance decisions by:

(1) showing relationships between result sets

(2) comparing result sets

(3) showing repetition in result sets

(4) allowing users to mark and annotate units of results at various levels of granularity

(5) automatic highlighting of query terms from the immediate query (as done in many systems) as well as from previous related queries as enable by search history

Saving relevance judgments along with documents is very important and is described in section 6.4.5.3.

5.4 Search task management

Managing the search task requires functions similar to those involved in managing other tasks: plan, review what was done, and evaluate. Search history can support many of these functions. The following sections describe search history uses linked to managing the task. Short, known-item finding tasks usually do not require extensive task management actions, as they are completed through a couple of simple steps in a brief time period. The tasks used in this study were complex topical ones, which can benefit from task management techniques.

5.4.1 Planning

Searchers usually do some planning before they start searching; pre-search planning is discussed in further sections. Several of the participants came to the sessions with notes handwritten on paper and used those while searching. It is important to provide support for systematically recording the planning notes in order to help users capture what it is they are looking for. Keeping this information on the screen while searching, or making it easily available to searchers, can help them focus on the task and evaluate search results more efficiently.

Since end users have less knowledge about the different information sources and search tactics, it is harder for them to plan actions. However, even though they cannot specify queries, sources or accurate steps, they have a generic idea about topics, keyword or names of people and organizations. Creating plans with even minimum information before starting the search helps with the management of the process.

Planning usually starts before the searcher logs into the search system. The information problem arises, usually along with some indication of the topic and other attributes. Depending on the context, searchers often create notes or otherwise document their planning. Catherine Best, at her website on legal research, suggests that searchers think about their topics and goals, and write down keywords and subject headings to search on, these can include synonyms and variations on words. During the planning phase expert searchers consider sources to check and directions in the search. Professional searchers would often familiarize themselves with a topic as an introduction to the search and collecting keywords. This may not be conducted in the search system, but serves as a preparation for the search and the results from it need to be used in searching.

The pre-search planning notes can be used to keep track of the search and check off sources or keywords. This function is discussed below in the section on Monitoring. This use of user notes is described below by a law librarian:

> Halvorson: How much planning do you do before logging on to QuickLaw?
>
> Best: I do a fair amount. I always make a list of keywords, synonyms and alternate words, and think about which words should be truncated. I think about what databases to search and make a note of them. Then, as I do the search, I refine it and check off what I've searched. Depending on how complex the search is and how many alternate terms it includes, my search plan is more like a diagram, with several columns of alternate terms separated by the appropriate connectors.

Planning is continuous all through the search session based on the constant monitoring and evaluation of the search process. Tools supporting planning, monitoring and evaluation can build on a record of user notes and search histories. Recording pre-searching and during-searching plans also provides better integration and flow between user tasks and does prevent an artificial differentiation between various phases of the search.

A part of planning or preparing for a search is to create a checklist of actions that need to be accomplished; this list can then be used as a guideline during searching. The search history combined with the pre-created checklist can tell the searcher what has been done and what needs to be done. The checklist can go through many changes based on what the user finds through the searches and if very informal.

I: Do you use it for anything else other than going back and forth?

S5: Just to remind myself of what I've already looked at.

5.4.2 Monitoring, keeping track

Showing the sequence between steps can also help users comprehend the relationships of actions and can promote system learning, as well as problem solving and planning.

While the searcher is executing her search actions, she needs to constantly monitor the inputs and outputs. Search history tools can help with this task, as they keep track of what has been done, and as indicated in the previous sections, they can also help with planning the future. Monitoring forms the basis of evaluation of the process, which in turn influences planning. Thus monitoring is invaluable for successful execution of searches.

One way to monitor the progress of a search is to follow the checklist created before the search started and mark the steps that were completed, and create reminders of new action items. Following the checklist guides the searchers' actions, helps him avoid repeating actions or leaving areas uncovered, and also helps with focusing on the main task.

Monitoring answers the questions What has been done? How good was the result? What needs to be done? In interrupted searching situations, this is a very important orienting factor for searchers.

Search histories automate this task to a great extent. By making the search history more easily available from the search screen, searchers may be more willing to use this record and thus plan and evaluate their search actions.

5.4.3 Evaluating action

Evaluation of actions happens after each step, when the user reviews and interprets results from queries, reads documents, etc. If he is not satisfied with the results, he takes remedial action. One of the law librarians interviews described how he encourages his students to constantly evaluate and check what they have done and found in order to verify that they are really clear on what they found and that they searched everything they planned on. This concern highlights the importance of recording actions and information for review by searchers:

Halvorson: What general warnings do you give your students?

Jackson: The most important thing is to be really sure what they've got when

> they find something. With electronic sources, when something pops up on your screen, it's easy to feel a sense of accomplishment or reward and think, "Oh, I got what I need." You might have gotten something, but you need to be sure what you've searched, and hence what you've gotten. You need to be sure you've searched everything you thought you searched. You have to observe the scope of the database. You need to know whether it's really full text or selected full text. You need to know the date range of a file's coverage. You need to know which source documents are included in the file. I see students do a lot of searches in the wrong database. They think they have included tax advice memoranda that were really in a different file. I see them not go back far enough in time.

Reviewing and evaluating previous steps can help users understand the current screen and how they got there. This is helpful in evaluation, planning and navigation (reducing disorientation).

During searching the user needs to constantly evaluate the results in order to see when to stop. This requires reviewing search steps, interpreting results, checking to see what actions have been completed and what have not. Jackson suggests that legal researchers make sure what has been searched to avoid mistakes. Often searchers assume they have covered an area when in fact they missed it.

> S5: I guess at this point I might stop and review my Word document and see if I have enough information to at least get started, to stop and read.

Constant evaluation of the results also help formulate the next steps of searching. The amount and quantity of the information, the presence of an adequate answer are all signs for searchers to stop. Stopping can happen for other reasons as well, such as budget constraints or interruption. This latest cause for stopping calls for the context recreating function described below. Stopping is often part of the search strategy, the user may stop search actions in order to review and evaluate previous steps and result in the search history.

Evaluation is closely linked to reviewing and interpretation. The extent of evaluation of actions depends on the task and the user. It can be fast or more involved. The system can help evaluation activities by showing repetition in result sets, clustering documents by user interest, and so on.

5.4.4 Creating reminders of what to do

Continuous monitoring and evaluation feeds back into planning through the creation of reminders for actions while conducting the search. Based on the plans and the history of the actual actions and outputs, the user continuously re-evaluates his situation and what needs to be done in order to reach his goals. In order to capture this rethinking of plans, the searcher needs to be able to change his checklists and other plan representations in the system. Creating reminders, thus updating the plan of action, is one way to do this. These finding will be built into future queries. Tools to tie the information found into future queries through planning and reminder functions are important.

5.4.5 Managing multiple tasks

Attorneys are often juggling multiple cases and may have to manage multiple searching tasks related to various legal issues. This may also be necessary for billing purposes. Time spent on different searches may need to be billed to different clients. Recording search history associated with different tasks can support this need. In order to accomplish this the user needs to identify the task he is working on to the system.

When saving results and search histories, it is important to note what task they were relevant for. Later the attorney may want to retrieve or organize them by the task they were carried out for.

5.4.6 Recreate context

When saving results or search histories, not only the search steps of resulting information should be saved, but also some of the context of when, how, why and by whom the information was saved. This contextual information will help the user reuse saved data. An important attribute to save is the purpose or relevance of information, as discussed in the search history data section. This is important for both the search history and the results, the user should be able to recall or recreate the context. Saving the system state is necessary for recreating the context described in the search history use chapter. Saving context preserves many of the elements that constitute system state.

In some cases searchers need to save sets of documents to be reviewed later due to time limitations or the pricing structure of the source used. It is important to save some of the context of how the searcher arrived at the set and what he thought about at the time. This will make further processing easier.

5.4.7 The need for recreating context

Context is operationally defined here as a representation of the state of the system and the user at a given point in the execution of a search or other task. The context is a time slice of the search history (only the last and several previous steps); the objects and actions represented on the screen and their relationships. In addition to the current state, recording the history a few states back is also a past of the context. The most important factor in recreating context is for users to understand what they had done and why, what they thought about what they saw and what they need to do next, presenting the history right before the system state can help with this, as shown in the next example:

S1: Yes, I'd like to know exactly where I left off. I guess I would certainly

> like to know what my search query was, I'd probably like to see the queries right before that, I would like to see what databases I looked at before that too. For example if I stopped and came back and I saw, ok I'm in fedall or all feds and I have this search query, I might look and see, gee, did I look at the 8th circuit first and then expanded it because I didn't find anything? I might look and see if I did a date narrowing at some point so that's still an option for me if I haven't done it. Ohm, I'd want to go back to at least three or four operations before where I left off, and that will help me reconstruct what I had done.

Personalization features are also very important, as the goal of saving the context is to put the searcher back into the same situation and mind set in order to help her continue exactly where she left off or to help her understand why a document was found and how the document should be interpreted.

The context recreated should help the searcher to understand:

! what he had done and why;

! what he had thought about what he saw;

! what he need to do next.

Recreating context is needed when the user must stop working on a task for one reason or another, and later has to return to it and start working on it again. In these cases, at the return the user needs to remember what he was doing and why, where he left off the task and what needs to be done next. He usually needs to review several recent steps in order to recreate the flow of steps and the reasons for taking them.

> I: And how do you use those? If you want to later use those little notepad files?
>
> P1: Then I go back and I'd copy the URL and I go to that. Or it would remind me that I found this out but I didn't find X, Y, Z out so I still need to go back and look at that rather than this.

Task interruption can be voluntary or involuntary. A user may have to stop working on a specific task because a colleague, a phone call or other external effect interrupts him. He can also decide that he will follow a tangent and then return to the main task line some time later having forgotten about what he has been doing. Even if the return is close in time, within the same session, the user needs reminders of what he has been doing, as forgetting is influenced by interference from search steps taken and documents seen while away from the task. Thus remembering is made more difficult by the actions carried out while away from the task.

In other cases, the searcher can save certain elements of a task for future follow up, in which case he needs to document what has been done so far and how to proceed to have a reminder when he returns. This saving for future review happens often when coming up to tangents in searching and deciding to follow the main task line, or when time limitation do not allow to examine promising but only marginally relevant results.

Attorneys are often working on the same case for a long time period, in which case they may have to stretch out their searching over some time as well.

> I: What would be that thing that it would be helpful for in the future? Like you would do similar searches?
>
> S5: If I wanted to search the same thing, or if I know I'm going to want to look at this again, but I don't have time to look at it right now. But I might have time a month from now.

In this case they will have to come back to the same search task group and continue on a different tangent; they will need a larger search context that shows several branches of the search and their status.

In collaborative team circumstances, a searcher may have to document the context of a search in order to pass it on to a colleague to continue the process. In this case it is important to record the preceding steps, current situations, plans and motivations for actions in a way that is understandable for another person, not just the creator of the record. This is also related to the scenario when the user has to recreate the context or review the search history to understand his steps in order to explain actions and decisions to others.

> S4: Oh no, I wouldn't recreate. just whatever is… , it's more just a history record. It could be that, if I have to justify a decision, which I haven't had to do but I could see it happen, something like, why did you give that opinion, and if it wasn't apparent from what I created, then I would have to do more search. I would have to try to recreate why I did it, why I said what I did.

Successive episodes of information seeking (Successive searching)

Information seeking sessions can be interrupted for various reasons (Lin 2000), and in some cases will not be continued. In other cases searchers will return to the task and continue it. For these cases, the section titled " Recreate context" discusses requirements for the search history interface to help searchers remind themselves of where they left off.

5.4.8 Information displayed to show context

In order to convey all this information and to enable the searcher to understand the recreated context, it is important not only to display the last screen and actions, but also to provide easy access to several previous steps, as described above. Showing several previous steps reminds the user of the process of actions. Searchers can more easily remember why certain steps were taken if the relationships between actions are clearly identified and they may remember what they thought of the results at the time, although this is better represented through user notes and annotations. In most cases the ultimate goal is to recreate the user's thinking, which is based on the steps he had taken and the results he had found.

> S1:It seemed like a nice idea at the time. Now I have to figure out what I was doing before I got distracted with the picture. Oh, I know, I was looking at the… is this that expert.. yes, this is the XY tire company , I was going to look at this and see, I know what I was thinking, if, as I suspect, a lot of cases here 118 cases, if they relate to prior testimony by expert and whether that's admissible in subsequent case, which I know this case touches upon, at least in part, and that's not what I'm interested in. If it is, what I'll probably do is exclude expert from my search query and hopefully narrow the number of cases. Ohm, so I was just looking this quickly to see where are my search terms. Actually that didn't really tell me that much about it at all. Ohm, I'll just check out the next case to see if…

Showing just part of the context may not be enough, exact depiction of user actions is important, as the smallest detail may be the differentiating factors that help the searcher remember her actions. For example, showing a screen of text from a document may not tell the user why she was looking at it, but scrolling to the right section of the text, highlighting the area read and pointing out the keyword searched on would give a much clearer picture of the process. This is also useful because it allows the searcher to continue the action right where she left off, instead of having to recreate the smaller details of the actions. In general it is important to follow the guideline outlined in the chapter describing search history data for all search history applications. (documents, databases, queries, etc.)

The participant in the next quote describes a need to highlight key terms or other text segments in documents that are viewed. Thus the initial processing of the document can be preserved for the next time the searcher needs to examine it. In the following quote, the searcher would like to see her previous query words highlighted, especially ones that she visited earlier:

> P1: I guess what I should have done when I find textbook, see I can't go back now and find it either. I guess it would be helpful to somehow be able to highlight terms once you found them so that you could go back to them and click on … if I didn't find another reference to textbooks, so I want to go back to that one that I did find and then maybe to click on those cases. But I don't see a way to do that.

As pointed out above, recreating context is not only important when the searcher returns after several hours, days or weeks, but is also important within the same session.

In saving search history information to recreate context, it is very important to allow the user to enter notes, annotations and highlights, as when returning to the task his own thinking may be the most important factor in helping him understand the process. Personalizing the history helps the queries and documents to the searcher's problems.

If the searchers started the information seeking by creating a plan, it is necessary to show progress against this plan. This may be very helpful in recreating context, as it does not only show the past, but also future steps. Integrating planning into search histories is especially important with this application, as the main purpose of recreating the context of an interrupted task is to continue the information seeking.

5.4.9 Keeping track of tangents

An important issue in task management is the question of tangents. A tangent is a line of activity not closely related to the main task, but emerging out of it. Information seeking (especially in hypertext environments representing highly interconnected literatures) is an inherently non-linear activity. Following tangents if often part of the search strategy through weak relevance, serendipity, and so on.

> S1: I'll probably pull up those quickly and see if there is anything interesting there. I'll just go through the last set of cites here and then I'll go back one set. The last group of cases is on bankruptcy so I'll ignore those. All right, so I'm going back to the latter part of the search where I saw a bunch of cases that were from my court and I'll just pull them up quickly and I'll see if there's anything there. I tried to start where I think I'm going to get the most immediate gratification and if I find something in that particular court that would be very helpful.

One form of potential distraction in legal research is the need to Shepardize cases. Attorneys often Shepardize documents while searching, which can distract them from the original goal and entice them into browsing through following citations from the original case. Tools to support searchers in returning to the original task line can be helpful with this task, as described in the next example:

> I: You mentioned that you would go multiple levels down and just Shepardize cases, would you ever want to go back to your original results?

> S1: Yes, almost always I do. Because typically, when I get sidetracked like that, I'm on case 8 out of 15. And so I'll Shepardize 8, 8 will be helpful, I'll Shepardize 8, I'll find another case that's helpful, I'll Shepardize something there, I'll find another case that's helpful, and then I want to go back to 9 out of 15. Once I've sort of explored that tree to its fullest extent, I want to go back to number 9, because usually at that point I haven't read it, sometimes I would write it down, 8 is very helpful and put a little star by it and go to no. 9 to 15, see if there is anything of interest there and then go back, but it depend on the amount of time that I had. If I'm really short on time, which I typically do, what I do is 8 looks great, Shepardize it, I'll print it, I'll look and see if there are any cases that cited that, that are either very good or very bad for me, and just use it, and maybe not even go through 9-15, depending on my time. If I had more time, I'd more likely just put down, I'd write down, maybe on a scrap of paper, you know 8, and then go thru the rest of them and then come

> back to it, so it depends, but in a perfect world, I'd always want to go back.

In a cognitive sense this is related to the management of the task, while the physical realization of this behavior is related to navigation. Following tangents involves the introduction of subgoals into the task, more or less related to the main task. These can be distractive, the searcher's attention shifts to a different issue and may forget to return the main task or when returning to the main task may have difficulties continuing it. (this latter case is described in the section on Recreating Context).

One of the challenges of tangents is to remember the main task line and to return to it. One way to remind the searcher that the direction he is pursuing is only peripheral to the problem is to attach some kind of reminder to the search history representation at the time of branching of actions. Another way is to create a reminder in a checklist format or a highly visible sticky note on the screen. In either case, the reminder should also take the user directly back to the branching screen so that continuation of the original task is easy.

> I: Do you think that that would be a different database?
>
> P2: Maybe not, maybe I should get back to the other one. I think maybe I should look at the other documents in the 35. Let's go back to that.

As discussed in the Recreate Context section, the more distracting the tangent is and the longer it takes, the harder it will be for the user to remember to go back to the main task and continue.

Planning, checklists for actions and search history records can keep searchers from going off on tangents or help them return to the main task if needed.

> S2: Upon reflection I really used my search results. A lot of times, it would keep me from going off on a tangent.

In the physical sense following tangents often leads users far away from the screen where the entry point to the tangent was or where they need to return to continue the main task line. It often involves navigating to a different part of the system, often following many links which causes problems with orientation in general and getting back to the tangent in particular. The physical issues of following tangents is described in the section on navigation.

Tangents need to be represented on the search history record, but should show that they are digressions from the main task line. After a tangent has been completed and proved pointless, the user may want to delete it from the search history record in an easy way. Another solution can be to keep the tangent but with a comment that explains its role.

As described earlier, often before a searcher starts following a tangent, she has a good idea of what she wants to do after she returns, but by the time she returns she may have forgotten it or parts of her plan. It is advisable to create a short reminder for after the tangent is completed in the search history at the place where the searcher needs to return to.

Showing the reason for going off on a tangent is helpful in search histories, it can inform the searcher about not only why he left but also what he was doing before he followed a tangent and what the natural next step is in the main task.

5.5 Search result and search history management

This section describes some of the current practices and recommendations for search results and search history management. Management activities encompass what happens between finding and using information, how the results from information seeking are interpreted and transferred to information use, and how search histories may help searchers during this task.

Search results are often used to inform the process of searching itself. Observation, interpretation, and judgment are part of the information seeking process, based on search results and other effects of actions. As the search goes on and after it is completed the searcher needs to interpret and organize the search results for use. A prerequisite for all of these uses is judging the relevance of the results.

5.5.1 Searching and browsing within sets of documents

The result set can be defined as all information returned to a query or clusters of documents or information found when browsing, e.g. all the cases linked from one case. The set can include the searcher's own documents (as discussed in section 6.4.4) and the cluster can be created by the user. In order to use the information found, searchers need tools to examine the result set. Providing a good overview of the information returned is very important, and so are functions that allow users to manipulate the result set to find out more about it and find specific information in it. Searching within a result set is a function participants often looked for. As described in the search history use section, a modification of the original query can achieve the same goals as searching within the result set, but often searchers do use this technique. For example, if a searcher entered a query: "copyright AND user interface" and later wants to find out if Microsoft is mentioned anywhere in the result set, participants would look for tools to search within the result set, as opposed to entering a new query: "copyright AND user interface AND Microsoft" and then look for the search terms. This may be a result of limited search expertise or experience or may be a solution to avoid many reformulated queries. If the searcher is interested in the appearance of several different companies in the same result set, the query to achieve the required action is rather complicated for an end user: "copyright AND user interface AND (Microsoft OR Apple OR IBM OR Xerox OR Palm Computing)". Executing multiple searches within the result set can be less complex and allows the searcher to focus the result set more. On more than one occasion, participants wanted to look for the query terms in the result set as a way of narrowing it down, as in the next quote:

P1: So it would be nice to be able to search through these cases for your terms.

Browsing within the result set often resulted in following inviting prospects outside of the result set and then returning to the set to examine the next item.

Even if the search system does not provide this function, the search history can support it as it contains specifications of the document sets. Manipulation tools for the data recorded in the search history can complement those of the search system. Searching across a selected set of documents should allow for the option of searching the user's own documents included with documents selected from a search system.

5.5.2 Search within document

Documents returned to queries are examined during searching to make an initial decision about their relevance. The depth of this examination varies based on time available, pricing structure of the database, the domain knowledge of the searcher and other factors. Searching within documents can help with this examination.

> S6: This is still that ALR article. So now I'd like to almost refine now. I see where it's putting my searches, the terms are highlighted, that's nice. Now I would almost like to refine "In this document" search. I would like to have a local search term within the actual document I'm dealing, trying to pinpoint. Because right now I want to find out more specific to toxic injury rather then just the tolling of the statute of limitations on a more general thing.

5.5.3 Compare results and document content

Repetition of actions may be a goal or may be something to avoid: comparing what has been done with what is being done can help with this. Comparing actions and result sets is an often-used tactic in searching, results gathering and information use. Independent of why the user wants to compare, he will need tools to support this task. As mentioned earlier, recorded information can be used to point out overlaps or repeating results or actions within the session. Showing overlaps between results sets, differentiating documents that have been returned before can help users in understanding the relationships between sets and collect results. It also helps them avoid repetitive actions, read or save the same case twice by accident.

Comparing actions such as queries or following links can help with keeping track of the steps of a search. These comparisons based on the various attributes of the actions, such as database, query terms and connectors, date and time of actions in case of queries. Presenting similar previous actions can also help remind the user of previous work she has completed in an area.

> S5: But since I was only researching one particular topic, I didn't think to keep writing everything down, since I was just focused on that one topic. If I was researching like three topics at one time, then I would probably write down a lot more, as I went along to keep …
>
> I: What would you write down?

> S5: I would write down exactly what I had searched and in what database I had searched, so that I didn't repeat myself, so that I didn't do the same thing twice. But that really wasn't an issue here, because I was only searching one particular topic.

Comparing results is one of the most important areas where comparisons take place. Often users want to know what percentage overlap is between two sets, what they have seen before from a set, which are the new documents and which are the repeating ones. From the repeating document, it should be easy for users to return to the previous query where that document appeared. This helps them evaluate the query and the result set, and find documents worth saving and processing. This kind of comparison between result sets can create displays where the new and the repetitive documents are differentiated by color and where users can easily generate groups and lists of new/old documents. In the following example a searcher would like to see which documents were relevant to both queries:

> P1: It would have been nice at some point, you know you asked about when search histories would be useful, I think it would be helpful to be able to compare search results from one search with search results from another. Especially since I have different issues, especially the school board issue and the separation of church and state issue, they are very closely related, but on the other hand they are not exactly the same. And some cases will appear in both I would assume and some cases won't.

Following the previous example, if a case appears in several searches it may mean that it is highly relevant, as it covers multiple research topics, as happened in the following example, where the attorney decided to save a case that repeatedly emerged in several results set:

> S6: I'm going to take this case because it seems to have appeared on the radar screen in a couple of areas I was interested in. I recognize it from… I saw it in a different section of the Keycites before and it seems to be, well I'm not going to read the whole thing right now, it seems to be a case that may be a …

Taking this kind of analysis a step further is to look at co-citation patterns, the similarity of citation patterns among documents, as the following participant described:

> P1: I know what you could do, you could compare these cases to the ones that came up when you put your terms in and see if any of them are the same because then at least you would know …
>
> I: And then which ones would you look at?
>
> P1: I'd look at the ones that both cited this one, because that cases is the closest I've come to and had terms like textbooks in them. Because I can't search in these for textbooks and I don't know if any of these were the ones that came up in the other one. Right?
>
> I guess you could just print them out.

> I: And compare on paper?
>
> P1: Yes.

As discussed earlier, the occurrence of the same cases over and over again can be a sign that the search can be concluded. When searching several sources, such as different databases, comparisons between the result sets can help searchers consolidate findings.

In the legal literature, documents can be repeated in different formats, as in the following example:

> S5: This is the same thing as the prior documents I've seen. This is obviously a pocket part to a treatise. So you are going to get a couple copies of that because different publishers put the same information out there.

Comparing the content of the documents is a more complex task. Marking the new information can cut down on examination time of the user. As mentioned earlier, comparing actions can help avoid unintended repetitions of actions and queries. Searchers often wrote down their queries in order to keep track of what variations they have tried before. By keeping track of all queries, the system can also identify frequently run queries or query clusters and make recommendations based on those.

Another method for facilitating the comparisons of queries is to highlight query terms from all previous searches within one session using a different color for each query. The strongest highlight or color for the current search, and highlight or colors fading out as the queries get older. This would serve the purpose of bringing interaction events closer to one another and would make relating results to the information-seeking task easier.

5.5.4 Saving and recording search results

Saving results is one of the most important features of search systems, as it provides the link between finding information and using it. This should be easy and flexible for users. From the user's point of view three levels of saving can be defined:

(1) Save and show

(2) Save and don't show

(3) Don't save

Elements of the search history that are saved can either be displayed to the user in a temporally ordered display, or can only be recalled when specifically requested by the user.

As suggested above, the search history use function defines how long search history information should be kept. Recorded information can be used to point out overlaps or repeating results or actions within the session, while cases that turned out to be irrelevant need not be saved for longer time periods. Other cases will be used in litigation and the searcher wants to come back to them and read or copy text when preparing a new document. Saving and recording may be needed for different kinds of searches. Short, fact-finding information seeking tasks and actions may not have to be saved for the future (although this is not true in all cases), while in the case of more involved and complex tasks the record can be used in later sessions.

Saving and organizing search histories can be very resource-intensive in terms of human attention. Automatic recording of search histories and system-provided organization of these records can help ease the burden of this task. The user interface should make it easy to stop recording the search history should and to delete items from it; this gives the user control over the automatic recording function, as described by the following participant:

> P1: But wouldn't you have …, if I could create it, if it didn't automatically create it for me, if I've said save this, this and this. I mean I make so many false starts and stops, if I could weed out the ones that, then that would be helpful, but if I saw all the garbage that I went through to get to that I wouldn't be interested in that. To me it would be more confusing for it to record everything. Because I'm not such a good searcher that I find everything. You know I find a lot of stuff that I have no interest in.

How it is done today

Searchers do save search results and search history now. They use many different methods in recording their steps and information returned to them. They take notes on paper, type in word processing packages, copy and paste text and URLs, and print screens and text. This points to the importance of studying the issue and designing good search history systems.

Emailing was another technique described to save search results. Emailing can facilitate later use and sharing the saved information with others:

> Halvorson: When you find a great page, what do you do with it? Are we just stuck with File Save As and remembering to save the graphics?
>
> Webber: There are three or four different things that you might want to do. […] You can send the link or the document by email. That can be helpful if you are doing research away from your own computer. You may be at a walk-up computer in a law library where bookmarking or saving would do no good, since the information would be stored on that computer and not be available to you later. You can email the page or the link to yourself at your email address and pick it up when you get back to your own computer. You also could send it to your client.

Printing as saving

Printing and managing paper records emerged in the observation and interviews as a heavily used method of saving and storing search history and result information. Printing has many advantages that paper documents in general have. It is easy to annotate and highlight, it is a tangible medium to share and transfer. Paper is easy to destroy; users are familiar with ways to destroy paper records.

> S2: At least before there was no way to save your Westlaw searches. So, by printing out your search history, your terms, connectors, and everything, instead of trying to recreate the wheel, you would know.

Printing is a widely used technique for saving documents and search histories currently. Search systems often pose limits on how long they will store a user's search history, the Westlaw system at the time of data collection stored search histories for two weeks. One participant mentioned previous experiences where the search histories were deleted at 2:00 AM each day. The printed copy is under user control, as opposed to the system-controlled electronic version of the history, the user can keep the printout as long as needed, can easily organize or discard it relying on well known techniques of paper manipulation.

> S2: at least before there was no way to save your Westlaw searches. So, by printing out your search history, your terms, connectors, and everything, instead of trying to recreate the wheel, you would know. And it was a lot easier; especially I used to highlight my search history, if you can… When you go back and you want to update you can actually see like the dates of the cases, and all that stuff. So it's a good … it's a roadmap.

Both documents and search histories are often printed to be saved. Printing can also serve the purposes of further processing, such as marking up search histories, as in the previous example. Documents are sometimes printed in order to facilitate reading. Several participants mentioned a personal preference for reading on paper, as opposed to reading on screen. A participant in the next quote expresses this preference:

> S5: Now generally what I would like to do is read it in hard copy form, so what I'm going to do is go through and print everything that appears to be explanatory of this law.

Marking up, highlighting and annotating documents for future use in the first step of information use. Current tools for online annotations and markings are rare and not very flexible, thus many searchers fall back on the printed option for these purposes. Annotated printouts can also be very easy to share,

> S2: I would highlight the cases that I actually pulled, the cases that I actually printed, highlight a note especially because a senior partner would look at your results. I would say, I used to search to see if certain insurance companies had filed certain products, and you know we know that this company is coming out with something and it's new, and I'd go and I'd do a search everyday to see, and I would print the results, highlight the name, wherever that company name showed up. And I would say not what we're looking for. So that would let the partner know, that you not only did the search but you looked to make sure that it was not ...

Copy and paste is often used to save selected sections of the text or smaller granules of search results than what the system allows users to save. Copying can also serve as a more flexible print tool, selectively save text sections, combine them in one document and print them at once. In other cases copying is used to save the full text of a document, if the copy tool provides and easier to use, more flexible solution to saving.

> S5: This is a segment, I want to print just this segment. So what I do is mark, copy, paste into Word and then save this document to my local drive. And I'll name it 'Klinger-Cohen' which is what I'm searching.

What is saved from the search history depends on many factors and is discussed in the chapter on search history data. Part of saving information is often organizing it in some way, this activity is discussed in the next section.

Whether notes, results and search histories are saved or not is influenced by how easy it is to save them. As expert searchers often take notes of steps and results, and suggest it to inexperienced information seekers, it is a behavior to encourage while searching. If saving search history information is easy, searchers may be more inclined to do it. Information in various formats requires different techniques for saving. Simplifying the process for all formats is important.

What format materials are saved in

What format to use is in saving search histories is an important issue as it influences storage and options for reuse. Thus, the format depends on the goal of usage, available technology, and personal preference and is also influenced by the individual ways users apply technology and tools to their problems. Saving is a very important function for the user, it is often not the first one to be implemented in the system and often implemented through a series of labor-intensive steps for the user.

In the next example the searcher decides on the format for saving a document based on how it will be used:

> Halvorson: Which format do you usually take?
>
> Best: It depends on what I want to do with the document. Sometimes I'm just printing it out and I'm not going to keep an electronic copy. For that, the HTML usually looks quite nice, and I'll just print it out. If it is something I want to save, and might want to use sections of in another document, I'll take it off in a word-processing format. Our office uses Word, so I would use something compatible with Word, like RTF. I have saved things in HTML and then blocked parts of them to use as internal quotations in documents.

While the goal is important in the next quote, the searcher also mentions context and available technologies:

> Halvorson: How do you handle data output from the Internet?
>
> Tyburski: It depends on where I am when I do the research, and what the attorney said he wanted. If I'm working in the office, I usually print it, hand the attorney a copy, and it's up to him to keep it. I don't save it electronically.

Flexibility of file formats is important in order to support the reuse of information and the integration of search into larger work processes. The searcher should be able to commit information to storage easily and later retrieve it in a variety of formats.

When searchers save information they need assurance that they will find it again later. One aspect of this problem is whether the information will be there or not. Search histories saved for a short time period may support some tasks, but not others, and not being able to find your history can be disconcerting, as shown here:

> S2: Well, before, I don't know how it is now, at 2:00 am in the morning your search … Like I would log out in the afternoon, it would say: "do you want save your search results", I would say yes. But at 2:00 am, they go bye-bye.

5.5.5 Deletion

The ability to delete goes hand in hand with saving, study participants expressed the need and concern for deleting saved information: search results, search history and personal notes. Deletion serves the purposes of housekeeping, it helps with task management and allows searchers privacy.

Archiving of search histories when they are automatically saved should be carried out by the system.

> I: when you take notes when you are searching, how do you use those notes later on?
>
> P2: I usually use them when I'm writing the brief I'll refer to the notes, but I won't keep my notes forever.

One reason for deleting search results and notes is that they represent an intermediary step to a final document, and will not be needed once the final document is created.

> I: Do you keep your printouts?
>
> S4: It depends. Rarely. If I went through my files now I probably wouldn't find a single printout.
>
> I: Why don't you keep them?
>
> S4: Because if I use it, it's for a memo and then once it's in the memo, then I'd destroy my…

In some cases materials and records can be confidential and may require deletion for privacy reasons.

When reviewing a search history, users should be able to delete parts of it. In the following example the searcher would like to delete tangents from his action history.

> P1: I make so many false starts and stops, if I could weed out the ones that, then that would be helpful, but if I saw all the garbage that I went through to get to that I wouldn't be interested in that. To me it would be more confusing

> for it to record everything. Because I'm not such a good searcher that I find everything. You know I find a lot of stuff that I have no interest in.

When users are responsible for deleting items in search histories, the system should provide tools to support this. Displays with easy viewing and deletion of sections of the search history or selected results items are needed.

The systems can provide support for pruning search histories by letting the user search on age, frequency of use, and incorporation into work products in combination with other criteria. The user can then decide which of the old information items to delete. Deleting results and search histories can occur at various times in the search process,

5.5.6 Results gathering

Gathering results from searching and browsing and preparing and personalizing these results serves as a bridge between finding and using information. It usually physically moves the information found from the search environment into the use environment. In traditional library settings the searcher would check the catalog (card or online), walk to the shelves and retrieve the document or request the document from a librarian, make photocopies or borrow the document, read it in the library and prepare notes or transport it to her office for reading and interpretation. Finally the notes or the new knowledge turn into a new document or new information through the use of the information found.

In the context of the current research, participants used their own computer for both finding and using the information. Tools to transport documents from the search environment into the use environment are present, but in their early development. To make up for this, searchers often used copy and paste to move information between applications. The download tools in the Westlaw interface version used were cumbersome to use, as were the printing functions. In order to print or download documents, users had to go through several screens, selecting one option on each screen, without sufficient feedback.

Results gathering is a very important function of search histories. The system can automatically gather results based on a log of user actions: queries and browsing create sets of documents, citation networks can also grow these sets, user action on documents can assign importance to documents.

Searchers gathered their resulting documents along topical and task-related clusters. Attorneys keep large amounts of paper files, including research files that contain records of searches executed and documents used. These are either organized by client ID or the research files by topical area. Often the documents would be preserved in two copies: one in the client file and another in the research files. This dual organization by task and topic can be easily preserved in computerized systems and also complemented by other attributes.

High recall is very important in legal research, as attorneys in litigation aim to find everything about a certain legal issue. Collecting results from various searches, sessions, and sources is important and search histories can help with this. Exhaustive and system-supported collection of documents is a function in demand. In the following two quotes a participant describes collecting documents in a search and the kinds of tools he would like to have to help this:

> I: I can see the point of having the shopping baskets at these sites where you can just collect things.
>
> S6: That would be very helpful. And then you can view your cart because at the end some of the things that I may download now become less important and I waste a lot of time going back to the download feature.
>
> S6: In the same way if you had a document management handler, say straight from, I guess I sort of do that, I could have been creating this system, but it would be nice if from Westlaw you would have your own set thing all ready to go, something like you create in your inbox or folders. Maybe you would get a questionnaire form at the beginning of your research where you would want to start setting up something like that, then you would be able to just click on the folder in the download section and just drop into there.
>
> I: So sort of a set of topics before you start.
>
> S6: Yes, possibly. It would also keep you an outline for your search so that you won't go off on some of tangents. I guess then you would do your search outline by how you set up your folders. Which may be helpful.

Two requirements for the user interface implied in these quotes are the following:

(1) Saving varying granularity of results, document collection to text sections. Search system should allow for many levels of the results to be saved.

(2) Saving results permanently or placing them in a buffer zone for later decisions.

This last requirement is described further in the next section.

Selective saving – organized collection tool

Another user need clearly identified from the interviews related to saving information is the ability to select information to be saved or select information to be deleted from saved records. Either the searcher does not use automatic recording, and the system only records and saves information specifically selected for this purpose by the user, or the system records everything and the user wishes to select information to be saved and information to be discarded. Information systems currently make "shopping carts" or other collection tools available for selecting and saving information.

> I: So when you are doing your little search histories in the Notepad, do you do the same thing?
>
> P1: Just write the one that gets me to all the other things.

The selective collection of results with the copy tool involves copying selected text sections or other document granules and copying them usually into a word processing document. At the end of the search the document can be reviewed, edited and finally saved and printed for use in other tasks. Copying in this case is used as in the previous case: collecting results in smaller-than-system-granules or in a more flexible way. The difference between this and the *Saving* use is that here the collection document can be reviewed and edited, it serves as a buffer between the search results returned by the system and the results saved and pre-processed for reuse by the searcher.

5.5.6.1 Integrate from multiple sources

Searchers often must integrate results from multiple sources. This situation did not arise in the sessions as participants were observed in one session only, when they did not have time to explore multiple sources. This situation, however, is one where search histories can be very beneficial in keeping track of what sources have been searched and collecting the results across many sources. A similar task is searching through meta-search engines, where the user enters one query and the systems sends it to multiple systems, presenting the results in an integrated way. Searchers often augment systems to keep track of multiple sources, which also shows the need for a tracking tool.

> S5: So now I have my Word document and I might go through it and highlight. So it looks like I printed from ten different sources, so this is just to give me an idea. I guess that's what I would do and now I would stop and review this, and then look at doing a more detailed search.

Related to this topic is the issue of identifying a core set of items from multiple searches. The search system can help with this by identifying results that appear in multiple queries, help users tag documents and show these tags in reappearing documents, etc. The shopping cart tool is also helpful and these functions can enhance this tool. This function could also help searchers determine when they have enough results or when they start seeing repetitive results and need to stop searching. (stopping and reviewing/evaluation discussed below)

When selecting documents for saving, comparing result sets and comparing individual documents can help searchers quickly see what it is that they have seen before. The topic of comparisons is discussed under the Search History Management heading.

5.5.6.2 Building personal collections

Attorneys build their own personal research files or collections from the results of their searching, as discussed earlier. As they usually work in one selected and often limited area of law, they can define the limits and outline of their area and build up the recorded research around these. They save pertinent documents returned to them by queries or found while browsing in the system. They organize and often annotate these and need to keep their files updated. If a topic comes up that they have worked with earlier, they may first look in their own research folder; they answer short questions from their personal research files instead of searching in a database. It is often quicker to go to their personal research files than going to a database. Their research files also reflect their personal history with the topic, it is in a way a personal database on the topic.

> I: How do you use this research file?
>
> S2: So the next time I'm doing, writing a variable life insurance prospectus, and this particular client wants to extend the … there is a certain period of time, free look period, and your state says that a 15-day period, the self-regulatory organization says it's a 15-day period. The life insurance company says I'll give them 30 days. Does that impact anything else? I'll go and look at my free look file.

Quick question answering is especially suited for the use of personal collections. It is usually much faster to look up a short answer in a local personalized collection than to go to a search system.

> I: So you would use your research files if you want to look up something and also if you want to do more searching on it you would go back to. Is there anything else you would use those research files for?
>
> S2: They are a good way to quickly answer question because if you think about it, a client asks you something, you do all the work, you put all of it in the client file, 10 years from now someone asks you the same question, are you going to remember the client?

These files exist on paper: creating a personalized collection should be supported in online systems as well. This can be realized through local copies of documents and queries linked to the database, or by creating a personal structure over the central database.

Documents created by searchers need to be integrated with documents written by other people or documents located through searches in information systems. Attorneys search for documents because they want to use them in their own documents, information found is reused in the user's writings, user notes are used in searches, etc. Tools to make transition between various documents easy are needed. Adding external documents to a personal collection that includes the searcher's own documents can help with creating links among original documents and derived documents, and flexible searching on combined results sets, and others.

> I: You were saying you wanted some kind of database capability on that?
>
> S6: I would like to have a database, have my whole system sort of database, like My Word databased, which would be great.
>
> I: So that you can search on different attributes?
>
> S6: Search for attributes, themes, terms, dates. I'd like to be able to search for documents that were created before a certain time that I could then take a look at and decide whether or not they are still relevant to clean it out. Unclutter

files.

I: What else? What other tools would you like to have?

S6: I would like to have something that would allow me to cut and paste in between the resources, easier to build … Building blocks for documents. Lawyers often use, reuse their work. When you build a memo on a subject, some of your information may still be relevant to a memo on another subject. Or it could be relevant to a memo on the same subject for a different client and you don't want to reinvent the wheel, it's a starting point, you may have to go back and check things for their current status and add things for the specific needs of the client, but you want to …, a lot of your value is the fact that you are supposedly an expert in an area and within that you need to write your expertise down. So you have the normal cut and paste, but there may be ways that you could have coded certain paragraphs and then you can build a document from an archive of coded typed resources.

I: So your own documents are part of that same collection.

S6: My documents and I would like to enter, do my own documents next to web resources, and this is one of the things that we are building for this EOS site, these are taking remote sensing data and bringing it to lawyers, and organizing it for the lawyers with the laws themselves. And so some of the things that I'm trying to do is to build a toolbox for lawyers to have data and law at their desktop.

Searching within the document emerged from the data as a needed function. Searching across many selected documents from a personal collection as a function would build on previously saved documents. The system should create links both ways, as discussed earlier.

Personal collections need to be kept up just as online databases. Documents need to be deleted, changed, reassigned, and so on. New topical categories are added as the area changes and the documents can be reorganized accordingly. A more detailed discussion on the maintenance of personal collections is in the "Search history management" section.

I: And if you get a new issue and you go back to your old files, do you update that? Do you go back to search?

S2: If there is something that comes up, big change of rule, used to be able to do this can't do this anymore, I go back to the old file, put a sheet of paper in there and say, in October 1996 the National Securities Market Improvement Act was adopted. And it affected suitability, state regulation of insurance products and leave the file as it was and then you start a new file. And then you also file separates, because then you need to go back and see. And every time an issue comes up that was affected by that you have to go back to the file to see how it was changed.

Local copies of documents saved in personal collections can be used to save annotations in.

5.5.7 Searching search histories

Searching the search history is an equally important feature. Searchers expressed a need for searching research files by multiple attributes in addition to topic and task: judge, court, client, date, title, defendant and legal treatment. This is more easily achieved in computerized search histories, if these attributes are saved with the results or search histories.

Searching within the history record is an important function to provide to users, as it enhances the usability of search histories. Many current history systems implement this function, such as the Internet Explorer history search. The need for search capabilities on the history is presented in the following quote where the study participant describes a searchable store of documents and favorites (bookmarks):

> I: You mentioned that you are building your own electronic library? What tools would you like to have in that, for organizing, or using those documents?
>
> S6: A way of having a searchable database of those documents. To be able to search them for subject matter. And word searching. Microsoft doesn't make it easy necessarily yet on organizing and using Access. Incorporating Access into a Word system or even in a favorites type of list. Making them, maybe have a few more questions when you are asking to put something …

In addition to searching by keywords or topics, participants often mentioned the need for searching by title, dates and other metadata recorded with search histories and results.

5.5.8 Organize search results and histories

Temporal order is the natural organizing principle for search history data and it is useful for presenting search maps. The advantages of this organization are many. It helps searchers understand the process, it supports the representation of the thought process of searchers. It is also useful in showing the relationships between events and actions by showing their temporal order. Although temporal presentation is useful, there are many situations when users must be able to organize search history data according to other attributes.

Search history must be saved in a form that allows for easy retrieval and, when applicable, for easy management and organization. If the search history information is hard to retrieve, browse and manipulate, it will not be used. When saving search histories, and results it is important to consider the different kinds of users and uses.

Participants saw organizing search history information and especially search results as a very important topic and discussed it in a large portion of the interviews. Searchers manipulate search history information through organizing it; they shape it more to their needs and liking; thus preparing it for retrieval and reuse.

Organizing can start much earlier than the search itself, it can start from the initiation of the project or the information seeking task, including the planning stages, as the major categories can take shape at this stage and need to be recorded.

> S6: In the same way if you had a document management handler, say straight from, I guess I sort of do that, I could have been creating this system, but it would be nice if from Westlaw you would have your own set thing all ready to go, something like you create in your inbox or folders. Maybe you would get a questionnaire form at the beginning of your research where you would want to start setting up something like that, then you would be able to just click on the folder in the download section and just drop into there.
>
> I: So sort of a set of topics before you start.
>
> S6: Yes, possibly. It would also keep you an outline for your search so that you won't go off on some of tangents. I guess then you would do your search outline by how you set up your folders. Which may be helpful.

The organization scheme can represent the final document written from the research findings:

> I: Do you file your notes so that, if the case file comes back, you could readily find your notes for it?
>
> Best: Yes. I keep a fairly structured file while I'm working on anything. [...] I've watched people struggle with being able to write, once they've done the research. They may get through the research okay, but then they have a pile of stuff they can't deal with for writing. They have to go back to the process of separating out the issues and figuring out what relates to what. The same case may relate to three discrete issues. You have to be organized mentally to get from a tangle of concepts to structured, well-reasoned analysis. Organizing your research papers helps that process.

Reorganizing is another way searchers manage their search histories, the original temporal order often may not satisfy needs of searchers. Reorganizing search histories and saved results is a way to do work, it reflects rethinking of areas, shifting priorities, growing understanding or changing tasks.

> I: How or why do you reorganize?
>
> S6: Because certain things are not necessarily … I make sort of like … if you picture the way a cache works on a computer, I create folders for things that I'm using in the near term. And then I make files for some of those things, or I may even take that whole near term folder and put it into another folder that's more of an archive. I may also find things are, sometimes I've had the habit of just saving things in the wrong place and not paying too much attention, so I'll go back and I'll look through documents that are in the wrong place and put them in the right place. I also find that some folders that I thought at one point would going to be filled, wind up having one or two documents, and then I want to get rid of those and put the information into categories that are more

likely be relevant.

Organizing is closely related to annotating, if something is saved and put in a certain place, there is a need for an annotation so that when the searcher comes back to it, it will be obvious why it was saved but without a good SH system in place sometimes the annotation is in paper form while the SH is on the computer.

5.5.9 Organizational schemes

The organization scheme can be developed by the searcher or taken from somewhere else. Classification systems of information resources can serve as one source of the organization scheme. The topic or subject descriptors assigned to documents or search histories are created by the searcher or taken by the searcher from the case text. These are usually highly personalized in the way they are used and are different for each individual depending on the kinds of topics they work on. The participant in the next quote described his taking subject headings from cases. Later the same individual elaborated on the fact that subject headings are not the same for different individuals.

> I: You organize the research file by subject, are those subjects that you come up with the definitions or are those that you take from somewhere else?
>
> S2: You take them from somewhere, but you know…
>
> I: You take it from the case?
>
> S2: Yes.

Various organizational principles can be used, the two most frequently reported schemes used in paper storage systems are topical and task-centered groupings. Topically organized stores are especially representative of results, but is also used with search histories.

Topical organization makes it easy to reuse information on tasks that concern similar topics. When an attorney is asked about a topic, she can first look her own research file, as in the first quote in section 5.5.6.2 Building personal collections.

The topical organization scheme can change with time and with changes in the interests of the user.

> I: So why did that change? You expected it to grow but it didn't grow? The folders you mentioned.

> S6: Some folders where my thought process might have been at one point, on a subject matter, I could be starting a project in an area and I create a subject matter called, say I'm doing a project on wetlands, I'm in environmental law. I'm doing a project on wetlands and I think there is going to be a lot of work that's needed to be done on habitat issues. And I find out as we go along in the process that habitat is not really the main thing, I think I may have files one thing under habitat, but it would be maybe more biodiversity in general what I'm looking for, so it's not just say species habitat, but it may also be plant and flora. Then I would want more broad categories
>
> because there is just too little in there to justify having a folder.

Attorneys will want to come back to the records and reorganize them, it is one way to do work:

> S6: Ok, so what I do is, often when I do searches is I take some hand notes, as well as I create folders and files on my computer to organize the information and the results of my searches. And then I go back and I reorganize, cut and paste.

The number of items in a category can influence whether a new category is created or not. For a small number of items the category is fine, but if there are too many items in it, it needs to be split.

Task-oriented organizational schemes are also used by attorneys to organize their search histories and results.

> S1: […] I guess what I might do is, when I save this stuff I usually save it as a Word document and what I might do is just create a subdirectory in my directory for this case where I have like my Word directories for this case, I'll have pleadings, correspondence, memos, if we're doing jury research I'll have a file for that, if I have expert materials that I want to keep separate from everything else, I'll have subfile for that. and I can imagine just creating another folder for miscellaneous research. And then as I go on I guess in the case, and I start to figure out whether or not these things are going to be useful I can even create subfolders for the research that I

Topical and task-oriented schemes may get mixed in attorneys' current organization schemes, as in the following example, where the participant starts describing topical organization, but changes it to task- or project-centered when giving the descriptions:

> I: How do you organize your own personal electronic library?
>
> S6: My own electronic library, the best example is, here are my favorites.
>
> I: Ok, your favorites list. So you organize them by topic?
>
> S6: I usually do by topic. Business resources, computer and software, demo is a project I'm working on. I usually keep, the projects have their own folder of favorites, like this is my ESIP. I have my own business, this is Earthpace, in Earthpace I have a whole bunch of folders and files related to projects I'm

> doing under that business, which is separate from the ESIP. And I keep a project on a convention, I have a project on a whole bunch of other things. And then I have about the environment in general. All kinds of different things. I'm interested in finance, I have a finance-government links and a whole bunch of links to various subjects, like where I live, Washington, travel, spirit, politics, news and sports, lawyer resources, art museums.

As discussed below, computer system allow both kinds of access simultaneously.

Physical organization structures: paper and electronic files

Attorneys and law librarians in the study reported on current storage and organizational practices for search histories, search results, and personal collections. The primary storage facility is paper file folders. A large amount of paper files can be accumulated in law offices; creating and maintaining these folders is an important task of law offices. The legal secretary and the law librarian are very valuable as they manage a lot this information. Paper-based documents are usually collected in hierarchically organized client and topic-centered research files. An alternate task-organized method was described by a law librarian, who organizes her notes and papers in an accordion folder related to a search.

In current paper-based research files search histories and results can only be assigned to one or two categories, attorneys described usually not making more than two copies for preservation in research files. Searching and retrieval is more difficult in paper filing systems, usually the only search tool available to the user is the hierarchical organization, although some of the participants reported creating cross references between folders.

Another currently used storage and organization method is the directory structure of the file system. Participants overall spent more time talking about these storage methods and expressed interest in more fully utilizing computer supported structures to store and retrieve search histories and results:

> S1: If I'm doing research and I come across a case which I think at some point may be one that I'm interested in, I may print it out or I may just create a file for miscellaneous research or something like that. I have not started to maintain electronic files of those kinds of cases, that's actually something I'd like to do…

Current physical organization structures, whether in paper file folders or computer file systems, are usually hierarchical. This may in part be due to the inherent structures of these media. Linear, temporal organization for search history information is natural in following the order of events and serve well for the user interface, but searchers should be allowed to reorganize this structure. In certain tasks a calendar can be used for temporal organization.

> I: And basically that's how you work, you reorganize, use. Let's see, how do you create and organize reminders, well you just said that.
>
> S6: Well, I do that and I also use a calendar.

> I: Even for your searching? Like you have to search on this and you will put it into your calendar.
>
> S6: Sometimes I may take notes in Outlook. I use notes in outlook, reminders, schedulers, put a deadline on a search, something, if I think of a keyword I will actually put it on my calendar, I'd say search for X, Y, or Z.

It is important to allow users to view their collection tool while searching in order to gauge how much information they have and how much more they need to collect.

> I: Is this a topic that just came up? That you haven't looked at before?
>
> S5: Right, I had no idea. All I knew that it was some law that had something to do with Department of Defense, and that's really all I knew. Periodically I save my Word document obviously, and I might look at periodically to see how long it is. So far I've accumulated three and a half pages.

Search histories include many different types of documents from a metadata standpoint: results found, user notes, documents written by the attorney. Search histories record many actions executed on these document, record their accessibility information, version information, user rights, etc. This information recorded in search histories can facilitate the management of these documents. The details of how search histories can support document management is described in the section on Search history and results management.

Tools to support flexible organization of search results and search histories are described in the interface design chapter.

5.5.10 Bookmarking

Bookmarking has been reported as an often-used technique by both attorneys and law librarians. Websites are bookmarked when they are perceived as being useful and with a potential for revisiting in the future. Some of the sites may be bookmarked because they are hard to find, or if a searcher feels she would never go back to site she did not bookmark and uses this tool as a reminder.

Some of the factors reported by participants included:

(1) difficulties arising from the locality of bookmarks to one computer when the user may be using multiple workstations;

(2) complexities of retrieving bookmarks when they are not organized stemming from multiple users using the same browser and bookmarks without setting up profiles or creating bookmark folders,

(3) the resource intensity of organizing bookmarks,

(4) inadequacy of the current bookmark management tools to handle a large number of bookmarks; over a certain amount of items it becomes more difficult to find a site among the bookmarks then on the Internet.

These reported problems all point to the necessity of better tools to organize bookmarks and other history items in general. The organization of bookmarks resembles that of search histories and results. Participants reported hierarchical and catalog-like organization schemes as well as a total lack of organization. Unorganized collections of bookmarks are hard to use and can encourage users to look for the information elsewhere. Topic- and task-centered organization structures are typical, with an intermingling between the two.

Some of the law librarians reported publishing their bookmarks on the web in order to share them with others, and many institutions formalized this practice by creating resource guides on their Intranets for employees. This again points to the benefit of sharing search histories and the tendency of searchers to record this information in order to share it with others. Another practice observed in interviews is the creation of a database of resources, including Web sites, for the use of a team or a whole organization.

Bookmarking is sometimes interpreted as tagging or marking documents or places as special. These marks are then collected into a list or a folder where the user can go to access these special marks. This may be useful for smaller granules than web pages, such as parts of pages, paragraphs and keywords. Bookmarking is marking at the web page level, lower level marking are also necessary in both the web and the database environment. Potential improvements to the bookmarking tool are described in the section on marking.

5.6 Information use

The ultimate goal of searching for information is solving a problem related to the information found. In current search system user support usually stops when the searcher locates the information. The continuity of actions provided by search histories can also help searchers in preparing information for use and making the transfer to using information.

5.6.1 Interpretation

Interpretation is a very important step in the process of finding information and also on the road from finding information to using it. It involves reading the documents found, often rephrasing them in some form, assigning meaning to them through linkages to current knowledge and recording the process and the results in order to build them into the knowledge network and mental model of the searcher. Interpretation leads to the development of mental models through the integration new information with the knowledge of users.

The process of interpretation and learning about a topical area is cumulative, past knowledge forms the basis of handling new information. Interpreting and integrating search results with old knowledge will eventually lead to the answer to the information problem; however without recording the results of integrating new knowledge into old models, the answer may be lost by the end of the process. Searchers can try to remember all the information found, but this may be difficult with only one exposure to it and lack of manipulation. Keeping track of this process in electronic environments through history-supported tools is an obvious application area of search histories. Recording results is a good foundation for recording their interpretation by the user and their linking to current knowledge.

> S2: Notes are crucial, because most people can't remember all the thought that went into their projects. You do research so that you can give somebody an answer.

On-the-fly interpretation is more typical of end-user searchers, as they are both more interested in and more familiar with the topical domain of the searches and the results. Participants often stopped searching for periods of reading, when they were skimming, reading and interpreting cases that they found. This is especially typical of domain expert searchers, as they are the ultimate information users and learn from the information as they go along. The process of reading while searching helps them learn more about the topic and thus form better queries or decide what the most appropriate next step is. In another searching task, it also helps with determining when it is time to stop searching. Reviewing accumulated interpretation records helps users decide whether they found an solution to their problem yet or not.

In a full-text database, relevance judgments are often based on examining sections of the text that are potentially highly relevant to the information problem (e.g., a high frequency of query terms) and thus potentially more accurate relevance judgments can be made. These examinations are also the first steps in interpreting the result and extracting information for further use. The information extracted is linked to the user's knowledge of the area and the problem updating the mental model the searcher has of a legal area. This activity is often accompanied by some kind of physical, behavioral actions: note-taking, annotations, copy and paste and so on (Marchionini 1995) in order to make the aggregation and use of results easier.

One of the consequences of the parallelization of search steps is on-the-fly interpretation of search results. Searchers enter a query, receive the result sets, start browsing the set and if they see relevant documents, they scan and read the text in more detail. After some reading and interpretation, they save the document or a section in some form, occasionally with annotation, highlights or other personalization techniques and go on to further queries, often strongly influenced by what they read and learned in the previous results. This on-the-fly interpretation process is usually accompanied by some kind of recording of the results and the user's interpretation of the results, including a external cues that relate to or simulate their mental model of the area and how their model is changed by the new information. Search histories can automatically record the results found and provide users with history-based tools to document their selections from the results, their thoughts in the form of notes, annotations and highlights and model representations based on these.

This interpretation activity in some ways makes up for the lack of searching expertise. Even though they cannot form perfect queries, they can perfect their exploratory strategies by probing the results sets more thoroughly and learning from them. Search histories are very important in this respect, as they can help searchers record their decisions and reasons for it and build on previous successes and failures.

> P1: Then I did another search using the new terms. I got better cases and so I got a couple of cases that I thought if I read them carefully I would be more knowledgeable and I could do even better searches.

Interpretation does not only help with understanding the topical area, but also supports learning about the systems and the information stored, forming mental models of these. By reviewing the kinds of information returned to a query and recording it over time, searchers make conclusions about the behavior of the systems and the characteristics of the information stored. Reviewing the query can help explain the results.

> S1: It occurs to me, just as I'm reading through this, I'll probably get, based on the searches I have, lot of cases like this where it talks about prior admissions or prior evidence or things like that, that would be either sufficient or insufficient depending on the case to establish summary judgment for one party or another and what I'm really interested in is trying to figure out whether or not a summary judgment ruling in and of itself is admissible, so now that I'm thinking this through a little bit more, I think what I want to do is go back and edit my query.

On-the-fly interpretation helps searchers by identifying the main issues of a topical area by simply reading the first couple of cases and then search on those issues further. Recording these issues from the first query in the planning functions of the search history tool and creating checklists can help with transferring previously found information into future tasks and thus shortening the cognitive distance between these events. Identifying issues is also helpful in building mental models of the area of the law.

> I: So you are learning about it now, are you finding out more things?

> S5: Little by little. It's obviously some kind of law that deals with government procurement. How some changes that have been made in federal procurement process, which isn't very exciting, but that does appear to be what it is.

Search history can help interpretation activities by keeping the searcher focused on the goal and constantly reminding her of the important issues and question to consider while reading.

Tracking the interpretation carried out by searchers is important in collaborative contexts as well. Often days or weeks after the search was completed, the users need to explain what happened, why he made certain decisions and what the results meant. Recording thinking is very important in this situation.

Recording the thinking of searchers is a more complex task than recording their action. Actions, such as open, save, edit query and so on can be automatically recorded by the system, but thinking can only be recorded if the user explicitly enters written/typed notes, annotations or voice recordings.

Systems should provide tools for entering notes and annotations linked to search histories. These help interpretation by allowing the user to reformulate what they found and link it to their current knowledge. In the words of one of the participants, the information found is a tool to forming an answer while what the searcher thinks about them is the answer. Thus the user's thinking should be recorded in order to be used in future work.

> I: So, that means that when you read something you want to reformulate it to make that answer?
>
> S2: Yes, you need to understand how that fits into what the question is and is that going to get me to an answer.

5.6.1.1 Interpretation in initial topical exploration of legal area

The type of the information-seeking task influences the steps users take in finding information. Even within the focused set of topically driven searches, user tasks varied along the dimension of user knowledge about the area: from initial exploration of a large area with issue identification to examining specific angles of a well-known topic. The initial exploration task usually involved topically driven browsing behaviors and a distinguishable set of search history use characteristics: While summaries of cases are helpful when the searcher has some knowledge about the area, they are less significant when the topic is brand new. In this case, the full text is more useful. When just exploring a topic with the goal of familiarizing oneself with the area, recording sources of information is less important than when the attorney will cite the source in a briefing.

As discussed above, in initial topical explorations, search histories can help users wander around in hypertext collections following topical links and collect relevant morsels of information (usually from the full text of documents) that provide insights into the nature and structure of the area. Search history tools can help users in the pre-search phase by recording potential keywords to be searched and then later to keep track of these and their variations added from results.

5.6.1.2 Mental model building

Sutton described the mental model building process of attorneys in information seeking. Attorneys make sense of and apply the new information found through building these models

Building a mental model of an area is tightly related to interpretation discussed previously. A mental model of the topic is in a sense the final outcome of the searching phase, the result of interpretation and the first step of using the information. By integrating it into the knowledge structures of the user, the new information becomes available for reuse in future work. Mental models are one way to think about knowledge representation in memory. Searching stems from an information problem, which means that there is a gap or other discrepancy in our knowledge (and knowledge representation) that we need to attend to. Interpretation in searching aims at applying new information found to the original state of our knowledge representation and fix the discrepancy. A record of what information the user encountered and how she reacted to it can help in externalizing our model of an area and thus updating our knowledge representation. Recording search histories and allowing users to manipulate them can help with:

(1) recording and protecting results of interpretation (discussed above),

(2) the heavy processing of information helps transfer it from working to LTM, and integrate it with pre-existing knowledge structures, thus building a mental model.

Tools for helping users to build mental models are discussed in section 7.3.2, such as concept mapping tools.

Users often mirror their knowledge structures in simplified physical formats, such as an outline of topical areas and issues in a document to be written or an organizational structure for storing documents. These external representations can be used as starting point in searching, they can be built into pre-search notes or shopping cart organizations and then later applied to searching, and refined during the search. Often search results and knowledge gained from reading them lead information seekers to rearrange these representations to reflect their new knowledge. Providing a tool based on earlier activities (activity histories) that are updated as the search progresses can help users refine their knowledge structures about an area. Structure should be complemented with notes, annotations, verbal explanations, and links to search results in order to better represent the knowledge of the user in a reusable format.

Often typed annotations and user notes serve as a behavioral counterpart to interpretation and mental model building. The notes can have a temporal role in helping the user form a correct mental model, by the end of which process they lose their significance and can be made inactive. In this sense flexible model building tools can support the interpretation steps of searching.

> P1: Typed annotations, I forget things all the time.
>
> I: But you wouldn't keep it, you would keep it for a session but not…
>
> P1: Yes, I wouldn't keep it, just the way I work is, by the end of the session, I'll cement it into my thinking hopefully.

Theories of knowledge representation in LTM have suggested many different ways of organizing knowledge. This suggests a potential need for very flexible tools for rearranging physical representations of knowledge structures.

Attorneys interviewed in the study developed elaborate paper-based research filing systems that they meticulously maintained and kept up-to-date. The organization structure of these files represented the structure of how they thought about legal areas in light of their practice area. The clusters changed time to time based on the tasks of the attorney and changes in the practice of the legal issue. This is a good indication of the usefulness of physically representing internal knowledge structures for document management purposes. However, this external representation can also be helpful in learning about an area through visual representation.

Assigning categories from the user's scheme to documents or document sections also serves as a kind of interpretation activity. Categories describe the user's knowledge of an area, usually the topical area of the information seeking. Selecting and assigning categories to results act as a kind of rephrasing of the information in the documents in terms of the pre-existing categories of the user describing the topical area or the problem/task.

Representing a previous version of the mental model allows the user to reflect on the change to it, thus reinforcing learning. Sharing mental models is important, as it provides a process of sharing information among team members. Mental model building is closely related to the search map.

5.6.2 Product/Document Writing

Legal proceedings involve the creation of many written documents. Attorneys use search results and their new knowledge to write new documents. Information use is often embodied in the writing of these new documents, although it starts during searching when attorneys learn from the results found. The documents, information, and knowledge gathered while searching have to transfer into the documents written by the attorney.

Writing often starts with saving quotes and other information from results through simply copying or copying with annotations. As in the example below, copying can help form the link between finding information and saving it in a form that leads to a document written by the searcher.

> I: Do you print whatever you read?
>
> S6: No, I'll take another first cut online, on the screen. And then what I often do is, when I have electronic versions, I will actually cut and paste from those electronic versions into one document, organized by headings, as if I was, you know bring the citations in there, I'll bring all the necessary stuff and then I'll actually sometimes use that starting to formulate my argument and I'll may then cut out quotes and use them as block quotes and then build up information around it and then take away parts of the case.
>
> I: so that's why you said you would need some kind of cut and paste facility between documents.
>
> S6: Yes, that would be helpful.

Copy from the search system and documents has been discussed so far. Searcher also copy from their own documents and enter the text into the search system or new other applications. As this type of copying often involves the reuse of the searcher's own documents, it is not considered here in more detail.

5.6.2.1 Collecting documents

Documents

Collecting documents in searching through results gathering tools supported by search histories can facilitate integrating result documents into new documents. Participants often arranged documents in clusters under various topical headings, which helped them make sense of the documents, and also monitor the progress of the search. These topical groups later form the structure of the writing or the organizational scheme of the attorney's research files. The organization scheme for arranging documents found overlaps with the organization of the document to be written in the following example:

> I: The clusters of documents or groups of documents were also ranked high. When you find cases, would you want to group them into your own categories and keep them like that?
>
> P2: Yes, I think so, particularly if I were writing a brief, I would probably, I might want to break it down into how I was going to divide up the brief and which case is going into what subject area, so that would be helpful.

5.6.2.2 Collecting pieces of documents

Information

As described in the last quote, participants used search results to cite sections of them. With searching and writing carried out in the same environment, it is easier to copy and paste between source and destination. Collecting text sections and inserting them into documents is an important functionality to support. One of the participants discussed his methods for copying and pasting:

> I: Do you ever use it for more searching, whatever you put into MS Word, do you reuse that later on in your searching?
>
> S3: like whatever I save? Well, no, what I'll do is I'll cut and paste from Lexis and I'll use that if I'm going to cite from that case, as the court stated blah blah, saves me to manually type it in there. Sometimes it also may be easier to email it to myself or I can email it to myself at home and read it print it out from word or my email software and read it at my leisure

The system can set up links automatically between the case where the cite came from and the document(s) where it is used. On the one hand, when quotes are copied from cases, the cases can 'remember' (documents or interfaces tool having memories or histories) that they were copied from, and the system can represent this information the next time the user looks at the same case. The case can display information regarding when and how it was used before, possibly with a link to the document where it was used. This can help create a continuity between work on similar legal issues and facilitate the reuse of the attorney's own work. On the other hand, the citation to the case can be automatically carried over to the document written by the attorney. This functionality is currently available through a Westlaw add-on to MS Word. On of the participants highlighted the importance of carrying the citation over to the document when writing:

> I: One thing I noticed is that you are not saving the citations of where that thing came from.
>
> S5: I would have to put that in, I would have to make sure that I highlight it, or write it in by hand. At this point I'm not writing a paper on it, I'm just appraising myself of what this law is, so at this point, for example, this is from the [blurred] Government Contract Litigation Reporter, if I was going to write a paper then I would need to cite it, I would definitely include that, but I'm just trying to come up to speed so I'm not concerned about that.

Keeping the links between the cites and the cited document including the source and the steps that led to the identification of the document can help any future work with the document. It can also be helpful in future citation checking tasks, in reusing the document or sections, in case the citations are questioned, etc. In other cases writing topically similar documents can start from an earlier document, in which case a quote with a citation included or a link from a quote to the query that brought it back can serve as the starting point for new research on the topic or for updating earlier research.

5.6.2.3 Collecting notes

Knowledge

Searching notes are good reviews of new knowledge learnt from the searches and can serve as the preliminary notes for a new document. In addition, to support their memory externally, searchers wrote down information in order to reformulate it and interpret it. In order to answer a question, this participant needs to reformulate information found, possibly in a written format:

> I: So, that means that when you read something you want to reformulate it to make that answer?
>
> S2: yeah, you need to understand how that fits into what the question is and is that going to get me to an answer.

User notes are usually the first version of the document, they need to be easily transferred to a word processor, even if they are not used word for word in writing.

> I: when you take notes when you are searching, how do you use those notes

> later on?
>
> P2: I usually use them when I'm writing the brief I'll refer to the notes, but I won't keep my notes forever.

Displaying the document draft while searching can also serve as a reminder or even a search action plan. Fast and easy transitions between searching and writing are important and are the link between information seeking and use. Computer displays can smooth this transition.

5.6.3 Reuse document building blocks

As attorneys specialize in areas of law, they become experts on certain topics. They often write about the same topic for different purposes, or in different documents. Reusing portions of documents that were already written on the same topic is a natural way to speed up work. The system can support this by examining previous documents and making suggestions for similar sections in both documents that came from external databases and the user's own documents. Integrating search results documents and documents written by the attorney can be supported.

> S6: I would like to have something that would allow me to cut and paste in between the resources, easier to build … Building blocks for documents. Lawyers often use, reuse their work. When you build a memo on a subject, some of your information may still be relevant to a memo on another subject. Or it could be relevant to a memo on the same subject for a different client and you don't want to reinvent the wheel, it's a starting point, you may have to go back and check things for their current status and add things for the specific needs of the client, but you want to …, a lot of your value is the fact that you are supposedly an expert in an area and within that you need to write your expertise down. So you have the normal cut and paste, but there may be ways that you could have coded certain paragraphs and then you can build a document from an archive of coded typed resources.

Reuse can be shown on the originating document, thus setting up links between tasks and documents that are topically related. By storing the user's own documents and external documents together, the system can analyze them for topical similarities and make suggestions. It would also facilitate future work, as it is always easier to start from older, already executed research on a topic, then without that basis. Organizing documents found in searches around tasks can support the integration of task execution in various applications by focusing attention on the task.

5.6.4 Version control

In legal information seeking, documents often occur in various versions. The versions of the individual documents and these relationships between various documents need to be communicated to the user through search results and search history displays. It is also important to keep track of version information of the searcher's own document. The history of a document can show what changes where made when.

5.7 Integrate with other tasks

Searching is only one task in the work of an attorney; integrating searching with other steps is very important. Since many of the tasks are now carried out using a computer, it is easier to help users transfer information from one state to next, from one application to the next, and from one task to the next. The system participants used did not support integration with other tasks; searchers and participants created individual solutions for bridging the gap between finding and using information. The interface tools proposed in Chapter 7 describe scratchpad and organized result collection tools that offer solutions to bringing information use closer to information seeking.

Creating functionality within the search system with strong integration into other applications where the saved information will be used should be a consideration in designing search systems. Providing interaction history information across various applications can also help with shortening the distance between applications and de-emphasizing the separation of user task environments in different computer applications.

5.8 Integration with other people: Coordinated (cooperative) work

Although collaboration is not at the focus of the dissertation, recording search histories is often required in coordinated work environments. Participants in the study often described the main reason for saving search histories to be sharing and collaboration. This theme was so strong in the data that it will be discussed here and can also for an avenue of future research. Often searchers record their search history in order to share it with someone else. The searchers in the study were employees of large law firms, where attorneys work in large teams with coordination, reporting, and task delegation needs. Tools to support sharing are natural extensions of search history tools.

Sharing search histories is also useful in intermediated searching. When a patron turns to a librarian for help in finding information, the librarian would like to find out what the patron has done already, what sources he had searched, and so on. In turn, when the librarian returns the information to the patron, he would include sources and some interpretation of the results, as described in the following quote:

> Halvorson: Do you have an intake procedure for new reference requests?
>
> Jackson: Whichever one of us is on the reference desk conducts a reference interview. I always ask what they think they are looking for. If they have a reference, I like to find out where they found it. I try to get a couple of

different access points of bibliographic information.

One important issue to consider in coordination contexts is that search histories recorded for the searcher's own use and for use by someone else may require different format and content. For use by someone else, more context and interpretation is necessary, usually in the form of annotations, written or spoken notes. It is also important to relate the history not only to the searcher's own task and context, but to the audience as well. This can cause problems if the originator and the audience of the search history have very different interpretations of the problem and the process. Another attorney might have different needs than a client or a librarian. While for an attorney a straight quote would be appropriate, an interpretive comment would serve the attorney's clients' needs better.

One way to share search histories and especially search results is to organize them by topical areas, which are usually fairly similar between attorneys on one team.

> S3: I would save the searches by area of the law of whatever issue I'm searching, that would be helpful to others because the attorneys in this office will come across the same issues.

5.8.1 Communicate and share search histories

As mentioned above, sharing search histories emerged as an important need from the interviews and observations. Many of the uses of search histories involve the communication of history data to others. Currently attorneys and law librarians share search histories and search results through email, printing and notes. The technology available to searchers influences whether and how searchers share information, as in the following example where a law librarian describes her method of delivery of search results:

> Halvorson: How do you capture and save information from the Internet?
>
> Botluk: It depends on whom I am doing it for. Often I just copy and paste into an email message, or just send them the uniform resource locator (URL) to look at. I usually do not save the research onto my hard drive.

Another way to share search history information in an intermediated setting is to publish complex queries for patrons to use. A variation of this occurs in intermediated search environments, where librarians or professional searchers carry out searches for end users; in this context it is important to record searches as patrons may return to retrieve the results again or with related information problems. This tool is used more and more in organizations in the form of request-tracking databases in libraries or information resources available on Intranets.

> Halvorson: Has the Internet had any effect on that initial stage?

> Best: No, but it will, once we get a significant number of research memoranda and other documents into an in-house electronic database. We are currently designing a research database that we can use through our firm's intranet. Once that is operational, one of the first things I would do is look there to see what we already have. Hopefully it will be user-friendly enough that the lawyers in the office will use it, too, and will look there before they even come to see me.

5.8.2 Teamwork, coordination

Participants of the study often worked in team structures and had to coordinate among team members through delegating tasks, reporting back on these including defending and justifying their work, sharing thinking and mental models. Two areas of support arose from the data in team coordination:

(1) Managing tasks among team members (delegation, reporting)

(2) Collective learning and decision making

If the team thinks it is necessary to continue with a search it does not have to be recreated. Because the decision making is delegated to a group of people, the context must be saved, since the individual does not know whether the search will be continued or not.

Searchers share search histories so that others can work with them: update them, reorganize them or interpret them. These uses require that other people can edit, change and annotate search histories, with the changes recorded.

> S2: And even with the search histories, I always print my search histories and I attach my search histories to a file, especially as an associate. A senior attorney can look at the same case and something totally different will pop out at them. And they'll say, did you look at this, and I'll say no, I didn't think it was relevant. Well it is.

5.8.3 Task delegation

Delegating or sharing tasks is an important function of teamwork. Multiple members can be working on the same task synchronously or asynchronously along temporal and spatial dimensions. This requires coordination between users and tasks. Tasks and subtasks need to be assigned, tracked, and reported upon. Delegation is often based on reporting on previous actions related to the task, and also results in reporting back to the delegating party. When delegating a task, one must include information about what has been done on the task so far and what needs to be done next, which is very similar information to that of the individual search history records for context recreation or planning.

> S4: It wouldn't be important to me, recording this whole session wouldn't be important for me, unless I was going to show it to somebody else for whatever reason, like to show them what searches I have already done and maybe if I'm

> working with someone else, I have done this search, so you don't need to repeat this, you can search other things.

One participant suggested including questions, notes, and an organized set of documents when passing an information-seeking task down to a junior attorney from a senior attorney. In this and other cases, team members build on each other's work and need a record in order to be able to continue with the next step, instead of having to recreate previous steps. This is also true when team members would like to go back and modify or fix something another member did. (e.g. modify a query they judge to be incorrect)

> I: So you are going to pass these full texts down to your associate to read?
>
> S6: I could.
>
> I: Are you going to put any notes on them, say this is …
>
> S6: I may go back and look through these and take a brief cut and try to now take what I took down and put them into some folders. Organize it by subject matter and make it easier and that way …

When associates report back to their supervisors, sharing the search history can be helpful in evaluating their work. In the following quote a supervisor describes how he uses search maps to supervise:

> S6: I would look at the terms they searched on, the database they went to first, I would look at what their movements were through the system, and I do that actually on one of the projects that I work on. I'm able to see the people who work with me on the project, what time they spent, where they went first, how they organized things in categories and who did what.

Keeping track of who is working on what task is helpful in avoiding duplication of work by different team members and allow them to take advantage of collective knowledge.

> I: You would give access to the other people to your search histories?
>
> S3: Yes, because you would share this info with other people, so we're on the same team, you wouldn't want them to reinvent the wheel. So they don't have to go through the same hoops you did. If you feel you have a good search, you can put that on there.

Delegation is also between librarians and patrons, corporate lawyers and law firms. In libraries between librarians, sometimes a different person has to take over a request for one reason or another.

5.8.4 Reporting and accountability

As mentioned above, a record of search histories can facilitate reporting back to the team, the supervisor or another team member. In team situations or in junior-senior partnerships, attorneys must report back about their work and the results of their work. This is sometimes done in order to show that the work has been done (e.g. for billing purposes). In other cases the results will be used in future work by someone else who needs to be informed of the results. A record of what has been done can support these tasks, emailing, printing and sharing in other ways should be easy. Annotation, notes and markings accompanying these can help interpreting other people's actions and thoughts.

The quality of work is often an issue in these reports as well, attorneys may have to go back and explain or defend a decision or a set of actions. Often this is done some time after the actual searches were carried out, so the searcher needs to record his decisions better in order to remind himself of why he has done certain things. More senior attorneys can use it as a training opportunity by correcting errors or by reinterpreting findings.

> S2: A lot of times you can see, as a young associate or junior associate, you go to the partner and they'll say why did you pull this case? And you'll have to explain what caused you to think that this was relevant. And it's difficult to do that if you looked at a hundred cases without a road map.

Often searchers need to see their exact path in order to reconstruct events and to be able to explain decisions. Using marking and annotations can help this, and the system can support it by showing everything the user did. In individual searching contexts the searcher communicates with herself across time, while in collaborative environments the searcher shares with others. Similar tools can be helpful in both cases. This can include actions like reviewing certain parts of the text, examining a group of documents from a large results set, etc.

Showing zero-hit queries has an important role in collaboration for showing that the work was done.

> S1: No documents, well that tells me something. it tells me that there's probably nothing that's directly on point in all federal jurisprudence at least since 1944.
>
> See, when I get something like this, one of the other things that I do file, I would print the screen.
>
> I: Print the query?
>
> S1: Actually just print this screen right here, cause it's got my query and it says all feds no documents satisfy… And I do that for two reasons. One that a week from now if I'm still looking for this I won't do this again, two, if the partner that I'm working for , who has a tendency to do this, says, there's got to be a case on this issue out there, I'll show him this piece of paper and say, back off pal.

Using search histories when reporting work completed can be helpful in sharing thinking and giving a more complete report of the tasks carried out. In the following quote, a supervisor would like to see his assistants' thought process through the search to better understand what happened and why decisions were made. In this case the report does not only support accountability, but allows the supervisor to evaluate actions:

> I: How would you record their thought process?
>
> S6: I would look at the terms they searched on, the database they went to first, I would look at what their movements were through the system, and I do that actually on one of the projects that I work on. I'm able to see the people who work with me on the project, what time they spent, where they went first, how they organized things in categories and who did what.
>
> I: How do you use that?
>
> S6: Not very much. I do the checks and balances when I question things, when I'm not happy with the way things have been put together.

Collaborative aspects of search history use require certain functions of the interface to be included and specific search history data to be recorded. The functions include sharing search histories, annotation and note taking, communicating with and informing other people, etc. A specific example of this is being able to share physical representations of mental models and road maps of searching. These functions and the necessary data are discussed further in the SH data and management sections.

Being able to account for or explain previous searching is especially important when there is a problem. If a decision is questioned, an attorney needs to go back and find out why he made it in order to protect himself. An extreme case of this is when attorney are involved in malpractice suits and need to provide evidence of their searching.

5.8.5 Collective learning and decision making

Search histories can be shared in order to facilitate collaborative learning and decision making. Often it is necessary for the whole team to be aware of new information discovered by one team member. Sharing search histories can help the responsible team member to explain and share ideas and findings.

5.8.6 Training

Using a record of previous searches is a good way to teach searching. In a team situation a senior member can share a query or a series of queries with junior colleagues in order to inform them about sources and search strategies. Librarians may use the same tactic to train attorneys. Search histories are also used to help diagnose existing problems in searches, and thus train through a history-supported help system.

> I: So here you can go back to Lexis, when you go back, would you like something there that's personalized?
>
> S3: [… If] you want to show a search that this is something you can do or this

> is a search I have done to help others learn, then yeah, I can see it would be helpful.

One of the responsibilities of law firm librarians is to keep attorneys informed about sources and legal information seeking in general. Creating checklists and pathfinders based on previous searches is an indirect way to use a record of previous searches in training and leading people in future searches. Publishing complex queries related to important areas of the law that the firm deals with is another way. Attorneys can run these queries periodically to keep themselves up-to-date on an area. A law librarian in the next quote describes the creation of checklists for specific types of information requests that librarians can use:

> Halvorson: Do you use a checklist to remind yourself of places to look?
>
> Chick: I've started to make checklists for company information and expert witness information. We get asked that kind of thing a lot and it's easy to forget a good source. I have a paraprofessional on staff and it probably would help him, too.

Senior attorneys checking junior attorneys' work also has a training effect.

5.8.7 Organizational memory

A database of previous searches and results can serve as the organization's memory in relation to information gathering. This is very important in large organizations or teams, where members may change over time, but their knowledge needs to be captured for future use on the project.

The form of this organizational memory can vary based on the tasks. In libraries the whole query can be saved or just the results or the sources. It can be published in the form of pathfinders or the whole database can be searchable. On legal teams, queries and results can be collected in one database.

Sometimes the organization or the state where the attorney practices may have a policy on how long and in what format records should be kept.

5.9 Administrative uses

Search histories can be used for administrative functions that require a record of employees' work. Billing is an important task in law firms that can take advantage of time measures of attorneys' work related to clients. Additional benefits of recording performance data is the capability to capture statistical information for evaluation, reporting, management and billing. There is a strong need for this type of information in organizations for management and reporting purposes. Automatic recording of the events creates an opportunity for creating customized reports based on the records. This can also be helpful with time management and efficiency evaluation techniques. Management often uses statistics to evaluate and report the organization's work.

5.10 Search history non-use

The main goals of the data analysis in the attorney observations and interviews section focused on identifying the user tasks where search history information can provide support, thus the presentation of the data in this chapter centered around the uses of search histories. It is important to discuss non-use of search histories and reasons for it. Several participants mentioned non-use of search histories while searching, as in the following example:

> S3: As far as legal research, as far as cases, I use Lexis-Nexis, I go to a specific database and there the search, recording the search is not that important to me, because I'll find whatever cases I need and I have the cites from there, the cites to those cases. I can always Shepardize them or research them further later, I don't necessarily need to record the search.

In this case, the searcher does not need a search history because he simply records pointers to the cases he found. Here, the searcher interpreted search history as a record of queries, not including results, and assumed recording results as a separate step. A similar occurrence happened in the next quote, where printing the text of the case is not interpreted as a search history function:

> I: Why not? Why don't you record search histories?
>
> S4: I just never had a need to. If I want a case I just print it.
>
> I: Do you keep you printouts?
>
> S4: It depends. Rarely. If I went through my files now I probably wouldn't find a single printout.
>
> I: Why don't you keep them?
>
> S4: Because if I use it, it's for a memo and then once it's in the memo, then I'd destroy my [records].

In addition to different interpretations of search histories, searchers with simple search tasks pointed out that search histories were not necessary for successful information seeking. The information problem is solved in a short period of time, thus the history record is not of immediate use to the searcher. In the following example, the searcher was pursuing one topic, had a simple task, and thus could remember all the necessary information:

> S5: But since I was only researching one particular topic, I didn't think to keep writing everything down, since I was just focused on that one topic. If I was researching like three topics at one time, then I would probably write down a lot more, as I went along to keep …

If the most important use of search histories in a given context is access to previous results, unlimited access to databases can make repeating a search cheaper and easier than managing history records.

5.11 Factors affecting search history use

There are many factors affecting the use of search histories, from individual differences to organizational context. Some of the most important areas are listed here. The study of individual differences in using search histories was not part of the current research, but is a very interesting topic for future research. In one instance during the study, two individuals in identical positions in the same organization presented markedly different use patterns with the suggested effect of personality difference, which led this researcher to an interest in the effect of individual personality differences. Job roles can also influence the use of search histories, attorneys and law librarians may have different tasks and needs that can be supported by history information.

Organizational context defines the cultures, rules and regulations guiding an employees' work life. It also defines many of the tasks of users. Information management is a part of attorneys responsibilities. They search for information, create, save and share documents over long periods of time. They need to reuse, store and find documents. Keeping search and other actions histories can help them achieve these goals.

Transaction-based pricing influences whether searchers want to revisit a document online in the system or save it in their own personal collection. If the pricing is a flat fee, it may be easier to save the customization on the system, or at least in relation to the system and not on the searcher's computer. In this case the searcher can recall the document from the system, as opposed to storing it in his collection.

> S3: When I did work for a law firm and cost was a factor, as far as how you searched, how you're charged, you were charged, at least for Lexis for the search, how many searches you conduct, you'd do one search and then narrow it down and you would be very careful to how many searches you did, and if I already did perform a search and was already charged for it would be good to have the search history, because you want to go back a lot of times you want to do the search again, the one that you did before. With unlimited access to Lexis it's not a big deal because I can just type in the search again, but with limited access it would be good to have search history to go back to a search you already conducted so that you are not charged again.

Another important factor is the pricing structure of the online database or other information source used. Transaction-based pricing influences whether you want to revisit something online in the system, or save it in you own personal collection, it influences how much on-the-fly interpretation is carried out. If the searcher is charged by the time she spends online, she will try to read less while online, and instead collect documents to be read later. The search history can play an especially important role here by noting why the document should be read. This can remind the person of the document better and place it into the context of the problem. The cost-benefit issue is decided by whether the price of access to the database is more expensive, or the user's time spent with saving documents and actions. Paying attention to saving too much information can lead to overload on the user's part.

6 Findings: Search history data

6.1 Introduction

Chapter 5 discussed search history use and, in this context, mentioned many types of search history data. This chapter deals with some of the same issues from a type-of-data perspective, but it does not cover all kinds of data. The process of information seeking involves many steps, actions, and objects. These have many different attributes (e.g. time of action, author of document) and have many different relationships between them. Of this multitude of information the most useful set should be selected for recording for each information-seeking situation defined by a combination of user, task, system, and context. What is this most useful set? According to Catherine Best, a complete legal search report should contain the following information about the search actions:

- headings and classification numbers used in searching digest and encyclopedia services, and the date or volume number of the most recent paper part checked
- headings used for searching in case reporter indices, and the volume number of the most recent paper part checked
- databases searched, date of search, and search string used
- case citations for cases reviewed, references to page numbers for helpful or damaging passages, and a record of where and when you noted up the case
- list of statutory provisions reviewed, together with date to which you checked whether statute had been amended, and a record of where you checked for judicial consideration of statute"
- title, author and year (or most current release date) for texts, citation for periodical articles, page references for helpful or damaging passages

(Best, 2001)

The last point describes data that should be included to record each reference found.

6.1.1 List of search history elements

Search history data can be entered into the system three different ways:

(1) Auto-recorded

(2) User selected

(3) User entered

The system can automatically record user and system actions, input and output information, without interference from the user. The searcher can enter additional information in the form of marks, highlights, notes and annotations. A third solution is when the system only records information specifically selected by the user. A combination of these techniques can satisfy most task contexts.

Search history information can be presented at varying levels of specificity. The following list of search history elements is used for the purposes of this dissertation:

- Session/housekeeping data
  - Project and client ID
  - Length of session
- Request data and query formulation data
  - Keywords
  - Queries
  - URLs
  - Name of database selected
- Search results data
  - Search results as a whole
  - User’s own documents
  - Individual document data
    - Author/person
    - Court of decision, level of court and jurisdiction
    - Relevance judgment
    - Topic of document and key number
    - Document text
    - Other attributes
    - Citations
    - Abstracts
- Marks and annotations
  - Mark
  - User notes
  - Annotate
- Links among documents
  - User-created links

6.1.2 Quantitative data on preferences for saving SH elements

Participants in the study were asked to rank the importance of recording 30 search history data elements for the current session, for several sessions in the future, and for forever. They were asked to mark the importance on a Likert-scale of 1-5. Understanding what "forever" means in this context proved to be difficult for searchers and required additional explanation. Participants were asked to fill out a form (Appendix D) after the initial discussion of search histories, searching the system and summarizing their searching. Follow-up questions were directed to them describing five of the 30 items on the list.

Table 6 presents the overall importance values across all three time periods for eight participants. All through this questionnaire, searchers made the assumption that documents are saved and described this assumption several times, which may explain why document is not in the first ten ranks. Attributes related to queries and results are intermixed in the set. It is important to note that "notes" made it into the first ten, yet notes are usually not recorded in current search history functions.

	History Item	Mean (SD)
1	Query statement	4.29 (0.65)
2	Relevance judgment	4.24 (0.73)
3	Name of database searched	4.22 (0.48)
4	Clusters of documents	4.19 (0.66)
5	Citation/reference to documents	4.14 (1.00)
6	Level of court of document	4.05 (0.92)
7	Notes	3.95 (0.30)
8	Jurisdiction	3.95 (0.60)
9	Type of search	3.95 (0.36)
10	Abstract/Synopsis of document	3.67 (0.46)

Table 6. Overall importance values. (n=8)

Rank in current session	**Current Session Item**	**Rank of item in several Sessions**	**Rank of item if kept forever**
1	Citation to document	4	-
2	Relevance judgment	5	6
3	Level of court	6	-
4	Clusters of doc-s	3	5
5	Query statement	1	3
6	Jurisdiction	7	7
7	Publication year	8	-
8	Notes	-	2
9	Name of database	2	1
10	Type of search	10	4

Table 7. Importance rankings by time period. (n=8)

Table 7 shows the importance of items for various time periods. This data is also presented in Table 8, where the shifts of importance between time periods is represented.

Query attributes:	**Current Session**	**Several Sessions**	**Forever**
Query statement	5 ⇑	1 ⇓	3
Name of database	9 ⇑	2 ⇑	1
Type of search	10 =	10 ⇑	4
Results attributes:			
Clusters of doc-s	4 ⇑	3 ⇓	5
Publication year	7 ⇓	8 ⇓	-
Relevance judgment	2 ⇓	5 ⇓	6

Table 8. Shifts in importance rankings across time periods.

Six attributes have been selected from the top ten list, three associated with queries (query statement, name of database, type of search) and three with results (clusters of documents, publication year, relevance judgment). Their position in the importance rankings was tracked among the three time periods in Table 8. Their increasing or decreasing rank in the next time cell is represented by an arrow in each cell. An interesting phenomenon was observed as the result of this analysis: the elements ranked very important for the current session are those describing documents, while as time passes, attributes of queries, reusable data element gained significance. While in the current sessions searchers would like to save attributes describing documents found, for future session reusable query elements are judged more important.

6.2 Session/housekeeping data

6.2.1 Project and client ID

Legal searchers organize their tasks around clients and legal cases. Searchers described their client and research folders as arranged both by client ID and topic. In order to support them in their task management, the SH should include a reference to the project, legal case or specific task. The participant below explains how she attaches the client ID to a search history she created:

> I: And what kind of thing do you put into that Word file?
>
> S5: Well, this one I just printed out, but what I might have done was gone to the top and typed in some thing, I might have typed in today's date, the client ID, whether this was an easy or difficult search engine to use.
>
> S5: Yes. And then there was something else one page back that I thought was important. [client ID] It was … just because you may need to associate, later on at a later time, you may need to associate this information with that file and just because there so many thousands of files, you are going to need to … . You don't want to find really good information and then sit racking your brain trying to remember what this was pertinent to. If it was a particular file, just because offices these days just have so many files and so many clients. You might as well make a direct association right from the start, maintain that. Of if it to pertains to ten different files, you would need to keep all ten numbers.

Recording the client ID is also helpful for statistical, billing and other administrative purposes in law firms. In libraries the reference staff needs to record the patron name for statistics, but also in order to be able to communicate with and fulfill their requests. This use of search histories is described by the attorney below:

> I: Ok, the client ID is very high.
>
> P2: I think that would be important for billing purposes, because depending on what billing system you had if you wanted to be able to justify the amount of time that you billed for the search. That would be good to keep a record of it.

6.2.2 Length of session

The length of the session is useful for billing purposes. Searchers are interested in the length of time spent on individual tasks or steps in order to improve their own efficiency or evaluate others. This information can also be helpful in understanding collaborators' actions, as in the following example:

> S6: I would look at the terms they searched on, the database they went to first, I would look at what their movements were through the system, and I do that actually on one of the projects that I work on. I'm able to see the people who work with me on the project, what time they spent, where they went first, how

they organized things in categories and who did what.

6.3 Request data and query formulation data

6.3.1 Requests

Law librarians interviewed emphasized the importance of recording requests in mediated searching environments. The original request and subsequent communications should be saved in order to best capture the user's need. Requesters may include some kind of search history with their request, or suggestions for the search. This can be captured and integrated into the search history. Examining what the librarian needs in this situation as a search history may be suggestive for search history systems and interfaces. Recording the request can also be useful in identifying reusable answers when a new, similar request comes in. Recording the date the patron needs the information by is also useful. Sometimes patrons come back at a later time and pose a follow up request or would like to get another copy of the answer.

6.3.2 Queries

Recording the queries is the simplest, minimal form of search histories. A query record should include the exact query formulation, with terms, connectors and field definitions, and values set for the results set display (e.g. ranking) and other output parameters. In addition, the name of the database or other source where the searcher ran or is planning to run the query should also be recorded. A query can be run at various times creating different result sets, it is important to record whether the query was run or not and if it was, when, where (scope of database or source) and what the results returned were. These attributes are needed for the interpretation of the results.

Recording result information along with queries is important; when recording queries it should be easy for users to access the result set from a query if it was run already. Alternatively, the query can be run again to provide updated results, see section 5.3.3.1.1. Reuse same query. Saving zero-hit queries is important to avoid replicating searches; they remind the user that even though no results were returned she has already executed the query. A lot of reformulated queries occurred in the study; recording the exact queries and showing the relationships among them proved very important. Reformulating queries has been discussed in section 5.3.3.1.2.

6.3.3 Name of database selected

The digital information stores are vast, finding the right neighborhood to search in is half the success. Legal information systems organize their information into collections that need to be searched depending on the general area of interest. The name of the database searched is important information to share for collaborative contexts as well. Retrieving search history information by the database searched can be important.

Pre-search strategizing often includes names of databases selected. User notes may have database names, these should be easily accessed from the user notes, and transferred from notes into action and checklists. It can serve as a reminder of what has been done, what needs to be done and provide easy access to useful resources. When creating reminders and checklists, database names identify tasks. In the following example a law librarian describes her planning and task management practices:

Database names are often reused. A special case of reuse is bookmarking databases that are often used by the searcher. These bookmarked sources can be a part of a customized system environment for the searcher.

6.3.4 Keywords

Keywords have many roles in information seeking, from formulating queries to relevance judgments and representing documents to be written, as described in Chapter 5. Saving keywords is important all through the search process, from the planning stages to the use of the information. Pre-search notes often contain keywords and terms to use in searching, recording these in the notes and then easily transferring them to queries and tasks is important.

If a user defines a set of interesting keywords in the beginning of the search, these may change during the process, new terms may appear and old ones get deleted. The changes in the focal keywords can reflect shifts in the user's interest, his thinking and actions, in his understanding and knowledge of the area. In the following example a senior attorney describes monitoring junior attorney's work, where a change in the keywords can reflect their thought process:

It is important to highlight keywords in result sets and queries to show the relevant sections. Highlighted keywords can help users focus on important parts of the document and find out quickly what the document is about. It can help searchers make relevance decisions and select parts to save, as in the following example:

> I: And then what else would you like to be able to do with that paragraph?
>
> S5: I might print it, I might associate it with the source, the title of the document, which is up here, so that I know where to go back to it. I make sure I got the authors' name, and the copyright information. I might go through the document and make sure I've got all the links. The way they set it up here, [the terms]. So I might go through and make sure I get all the instances where it happens, where the phrase appears. If it appears throughout the entire document, I just go ahead and print the whole document. This is 29 pages, so I might print that whole document if I found that it was pertinent in its entirety rather than breaking in up just tone in on the words around what I was searching.

Highlighted keywords from previous queries are also important to display, as this ties in the various steps of the search process together. The keywords that highlight in a case depends on where in the process the cases is, thus the same case can have different sets of highlighted keywords at various points.

6.4 Search results data

The goal of searching is finding information, and often only the results are saved and carried on to the next step. Search results are the most likely candidates for reuse, in one form or another. Searchers used many ways to save results as described in Chapter 5: printing, saving files, handwritten notes, emailing. Not all findings are worth saving; tools for selectively collecting and saving search results are necessary. Flexible collection tools ensure the bridge between searching for and using information. These are described in Chapter 7: Interface design.

In legal research results are usually documents: court opinions, law review articles, news and other types of articles. Searchers in most cases need the full text of documents, although the flexibility to only save a section of the document is needed. The following sections set out to answer the question of what search results data to record.

6.4.1 Search result set data

Save number of results

The number of results returned to a query is very useful for certain tasks. However, this number changes over time as discussed earlier, these changes should be communicated to the searcher. Queries with too large or too small result sets are more memorable to users, as they stand out and in some way they shatter the expectations of the searcher. This is a characteristic that can be recognized by the system. The number of results returned to a query also helps users learn about the system, find out what the collections strengths and weaknesses are, how documents are indexed, etc. It can also be a signal of a problem with a query.

Another important use for this number is to orient users within results sets and search histories. Telling users that a case is no. 33 out of 100 cases returned to query no. 5 out 23 queries will place them exactly at a point in the search process. This reduces their confusion and helps them understand their current position in reference to the whole.

Preserving link to source when saving

Participants were observed to save results including their source and the query that brought them back. When saving local copies of results, the link to their original source should be preserved. If only a quote from a case is saved, users should be able to go back to the full text. It is also important to link search results to the task they are relevant for. After a while a user may forget what topic was relevant for what case, preserving these relationships may be helpful.

Save actions executed on results

Users want to manipulate saved search results, and capture the outcome of these actions. Next time the users encounters the same document a track of these changes should be reflected. The kind of manipulation tools needed will be discussed in the next section on search history management. When saving results, it may be important to show what was done on which documents, which documents were opened, read, etc. as this is easy to forget.

In intermediated contexts and coordinated tasks, finding search results for others and sharing them is the ultimate goal. Search intermediaries will not use the results in any way (other than sharing and maybe storing them), and in a collaborative environment a staff member may only carry out one step of the search (e.g. updating by paralegals). In these cases, search results are managed only in order to be shared and/or preserved for others. Preserving this information is a way of managing the searchers' knowledge in order to reuse it in future steps.

When search results are used by someone other than the person who found and saved them, different information should be saved, often more information, as the other person has no knowledge of the situation and has to learn it from the search history or results record, or from additional, possibly verbal, information from the creator of the record. Saving context is helpful in this case, a description of the issues to be considered when saving the context is described in section 5.4.8. The patron may need the search result in a specific format.

6.4.2 Individual document data. Descriptive.

Participants mentioned a number of document attributes important to them. Some of these are standard metadata elements; others are specific to the legal domain. This section brings together user responses on this issue. Some of the attributes come from the information source the user is searching, while others can be generated locally by the search history system or the user through notes, annotations, and other user inputs. While these attributes are discussed separately, they are treated together as a document is processed.

Saving data about the individual document is important for many different reasons. If a user saves a document in full or in part, we may assume that he finds it important. Individual documents are identified by a set of metadata that are described in the sections below. In some cases users do not save documents that they later need (delayed relevance). The system needs to keep track of pointers to all documents that appeared in the search, as later the user may want to go back to them even if he did not know that at the time of first interacting with the document.

6.4.2.1 Citation data

Citations in legal literature are used to denote court opinions, journal articles and other documents by a shorthand created according to a common set of rules. A citation to a case consist of three or four elements:

(1) The case name

(2) Its locations in the legal reporters that collect cases

(3) Court and date of decision

(4) Information about the authoritativeness of the case, where needed

The information represented in citations can be parsed and used in other displays or visualizations. Breaking up the citation for user interface purposes may be appropriate in some scenarios, but not in others. Experts are used to reading citations and do not need field labels.

These shorthand references are widely used in the legal literature for pointing to cases. Saving these references can help with storage issues, as saving the citation enables the system to later identify and recall a document, without having to store the full text of the document.

> I: What’s in those notes? Are they citations or keywords, or summaries?
>
> P2: Usually citations and summaries of the cases, but I think they might be better way of doing than abstracts.

Users will need to insert citations in their text when referencing a document. Automatic saving of these citations by the system eliminates the need for the user to handwrite citations into her search record. In the following quote the user describes sometimes handwriting citations into her search records, automatic saving can save time for the searchers:

> I: One thing I noticed is that you are not saving the citations of where that thing came from.
>
> S5: I would have to put that in, I would have to make sure that I highlight it, or write it in by hand. At this point I’m not writing a paper on it, I’m just appraising myself of what this law is, so at this point, for example, this is from the Government Contract Litigation Reporter, if I was going to write a paper then I would need to cite it, I would definitely include that, but I’m just trying to come up to speed so I’m not concerned about that.

Citations are important components of search strategies. Searchers may start searching around a case that is particularly on point or answers their needs. Creating a list of citations for future updating is another way to use citations in search strategies.

> I: Tell me about your notes that you started with.
>
> S6: These are from a conversation with a co-counsel who actually had some citations that were written down for us to check out. They were also some

> notes that I took on terms that I thought I was going to want to search. Keyword, as I was thinking about this how I frame some of the things that we were going to do.

6.4.2.2 Title

The title of a document is a high-level surrogate representing the document in a compressed format. It expresses the content of the document; the title of a legal case identifies it through the names of the parties, such as in Smith vs Brown. Saving the title in addition to the citation is important as the name of the parties can remind the searcher of the facts of the case and how it related to their problem.

Titles are included in summary representations, and with legal cases they include the names of parties as well. Users sometimes give new titles or short description to saved documents that are more meaningful for them.

6.4.2.3 Author/Person

There are many people (and organizations) involved in legal cases and other documents. The author in legal information varies depending on the type of document. Participants in the study expressed interest in various human roles: attorney, judge, document author, parties involved in case. They wanted to record and search by these roles. When recording various persons and organizations, their roles should be recorded as well. These attributes should be included in document surrogates, as they affect the relevance of the document.

> S1: That's great. OK, put this down, this is good development. Ahmm, particularly if you happen to be doing some research and you know that there are some cases sort of in the general subject matter that you're interested in, or if there's a particular judge or court that you're interested in you can very quickly, that's why sometimes I'll start with the cites and if I see something, for example, if I see something that is an opinion by my judge, I'll usually start there. And even if it is a case that factually isn't similar to what I'm looking for, at least I'll be able to figure out the kinds of, sort of brush stroke cases that this particular judge is interested in. And like for example, even with something like a summary judgment motion or something like that, where you have a number of very broad principles, you cite to the federal rules of civil procedures, you cite to a handful of like really broad cases that deal with summary judgment, I'll sometimes take the language right out of that judge's opinion, or cite the same case at least, so at least the judge will look at my case and say, that looks really familiar to me, and it's just a good way to start. But I don't have the luxury here, because I'm looking for the needle in the haystack, so unless I get lucky… I'm going to go and see the cite list real quick first and see if there are any cases that I'm …

6.4.2.4 Court of decision, level of court and jurisdiction

Precedence is cornerstone of the American legal system. The law is defined in court rooms and published in the text of cases. The level of the court and its jurisdiction define the authority of legal cases and rulings. This makes the jurisdiction and the level of court very important attributes of legal cases to save when saving these documents. These elements are usually represented in the citation to the case, so that is a good surrogate for the case itself. Searching and sorting by court is a feature of legal databases and should be preserved for the search history. It is also important to include this information in abbreviated representations of cases, as the court can decide the relevance of the case.

> I: My other question was you said if you have easy access to these things. How would you like to find them later? So if you had a big database of everything you found before, would you go by judges name or fact situation of what would you go by?
>
> S1: I would probably organize cases by subject matter, and then I guess what I would like to be able to do, it depends on the volume of the cases that you have, for example if I had in Word, you know, a folder for miscellaneous research and then a subfolder for that on, expert testimony, basic case law on expert testimony and let's say I'd have 3 or 4 cases in there, it really wouldn't matter to me if I could search that, I'd probably just flip thru all of them pretty quickly. If I had 50 cases in there, it'd be great if that was searchable so that I can quickly determine the court, the year, the judge, the parties, maybe what the basic holding would be, or something like that. And I guess if there were a way, the first time I pull up a case I could just sort of verbally dictate what that is, then have a link to my own voice, which would drive me a little crazy, but would be useful, I can imagine that would be something very helpful.

6.4.2.5 Other descriptive attributes

Other attributes such as publisher or copyright information can be important in some cases and are usually included in summaries or representations of cases. The language of the document is English in most cases and does not need to be preserved separately. The length of the document is important for users, as they would like to know how much time it will take to read, how many issues were covered, etc. The length of a document gives an overview of a kind and also orients them once they are inside a document.

The publication date of the document is again an important piece of metadata, in precedent-based legal systems it is even more important, as the date of the decision can be crucial in judging the relevance of it. Date can be a searching and sorting attribute in search histories just as in bibliographic of full-text information systems. For statues the dates are important in considering amendments. Date of publication and decision are the same for legal cases.

6.4.3 Individual document data. Content.

6.4.3.1 Topic of document and key numbers

Topic and subtopics of a document are very important in making relevance judgments and also in reusing the document. Saving these is an obvious need, extracting and representing the topic in a succinct form is a complex task. In the Westlaw system key numbers comprising of classification numbers and terms are assigned to the document; the search history system can take advantage of these. In other cases however, it is harder to extract the topic description. In some cases the topical aspects interesting to the user are not included in the system-assigned descriptions. (Once a user has made a few assignments to a user-defined topic, the system could use these to train an automatic classifier which then make suggestions.) When the user organizes documents at the time of saving, this often takes care of representing the topic of interest by assigning cases to certain clusters, as described in section 6.4.2.

Topics are often included even before the search starts in the form of keywords in user notes. These can be carried on to the search and represented in the search history as an organizing principle, as they represent the different topical aspect:

> I: Yes, it's usually not continuous text, it's keywords, reminders.
>
> S6: yes, I usually go into searches with keywords. That's the one thing I try to think of because searches are all keyword, or mostly keyword-organized. Phrases, of concepts, sometimes citations, but it's really the keywords.

The same is also illustrated in the following quote:

> I: how do you organize your research file?
>
> S2: By subject matter. So I would have variable life insurance and under variable life insurance I would have everything from prospectuses for public offerings, private offerings, I'd have files for tables, graphs, files on language to explain variable life insurance, you have a research file room.

The topical clustering of documents is very personalized:

> I: when there are two attorneys working on the same case, do you think their research files, their subject groupings would look the same?
>
> S2: no. Different people have a need for … people's needs for information differ. I like paper. So in my research files there is a file on how the NASD looks at suitability, and there is a suitability file for states. There is a suitability file for variably life insurance products, and there is a suitability file for variable annuity products… you know, because that's how I… When I had an issue, I wanted to be able to go to THE [participant emphasis] file with THAT [participant emphasis] issue, when another associate would have a suitability file, so theirs should be this thick. I'd have ten suitability files, so I want to space it different
>
> I: because you're working on that particular issue.

> S2: Yes, I want mine to be what is the regulatory agency name, what is the state name, and they just want to know suitability. They'll pick through all that stuff to see which state.

Topic should be part of the document representation, as it makes it easier for searchers to remember the document than just simply by title. When they want to find the document again, they may remember the topic, but not other attributes as in the following example:

> I: So you would use your research files if you want to look up something and also if you want to do more searching on it you would go back to. Is there anything else you would use those research files for?
>
> S2: They are a good way to quickly answer question because if you think about it, a client asks you something, you do all the work, you put all of it in the client file, 10 years from now someone asks you the same question, are you going to remember the client?
>
> I: So you don't want to look it up by the client, you want to look it by the subject.
>
> S2: Yes, with some insurance companies, or with some crook…
>
> I: So you remember it better by the subject?
>
> S2: yeah, but even if you don't remember it, go to your research files list, and see if I have a file of it. And it's just like you said, I think I may have a file of it.

Topic detection may link current documents to previously handled ones. Classification numbers, key numbers in the West system can help with this function. The use of the Westlaw key numbers is largely influenced by personal preference. They represent topics according to the west indexers, and not to the user's needs. In some cases they may be useful to save with cases, as they can be reused for further searching.

6.4.3.2 Events of the case, legal issues and holdings

The events of the case can be easier to remember than the ruling or the legal issues. If the case has a summary, that is a good candidate to represent the case in aggregations, but when there is no summary, the events of the case are difficult to extract automatically. Case events thus can remind users of the case and bring it back to his memory. However, the legal issues and holdings of the cases are very important in the reuse of the information, and as they are harder to remember, they need to be represented right along with the events of the case.

The parties in the case are also interesting metadata and often important for the case. The parties are usually included in the title of summaries if the events of the case.

6.4.3.3 Abstracts

An abstract is a brief summary of a document, representing the important points of the content. Court opinions in the Westlaw system include an abstract called synopsis (Figure 6) at the beginning of the case text summarizing the main issues and facts of the case, this is called the synopsis. In addition to this overall abstract, there are abstracts called headnotes summarizing the different legal issues of the case, with links to key numbers, the Westlaw classification numbers for the legal issue. For examples of a synopsis and a headnote please see Figure 6.

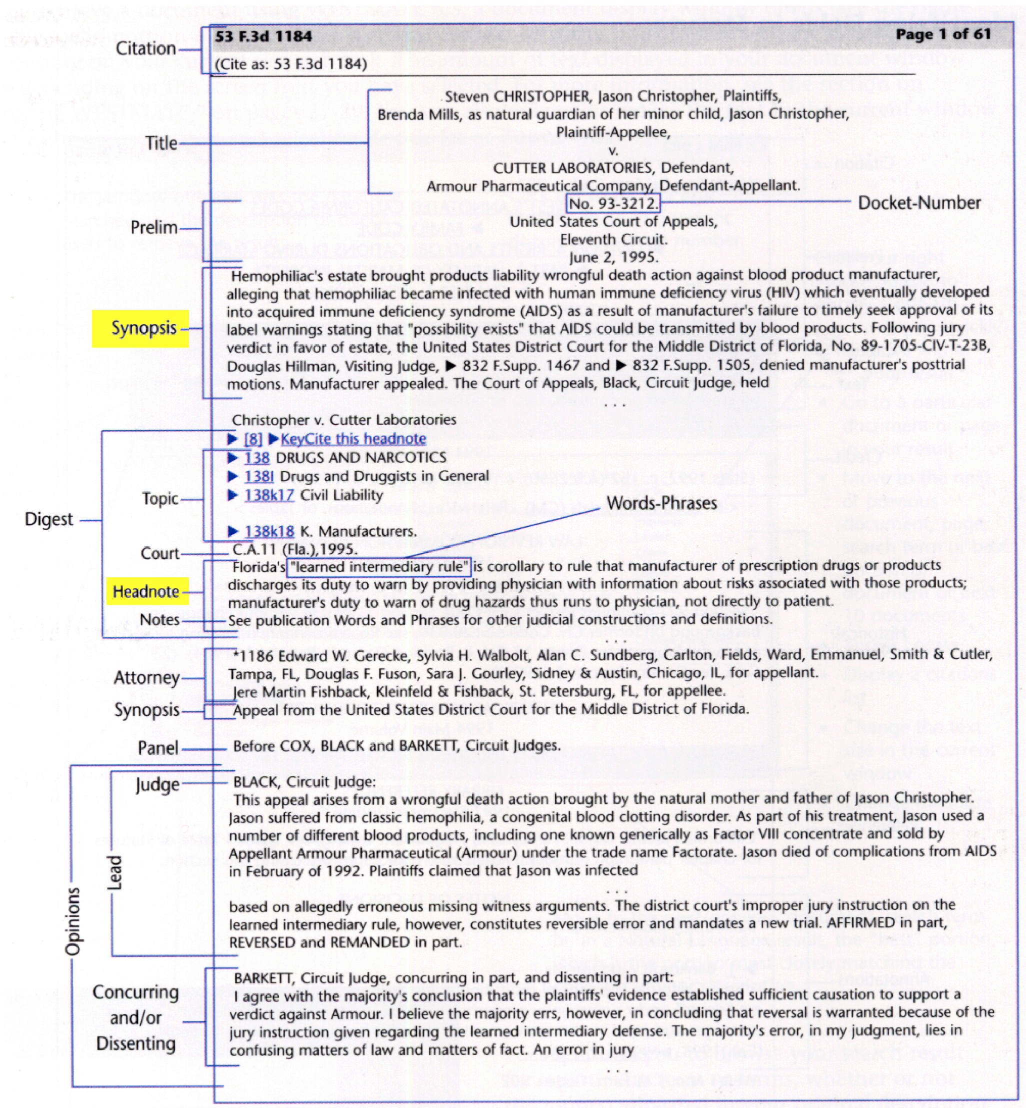


Figure 6. Sample page from Westlaw court opinions with fields.

Headnotes represent an important added value feature in the Westlaw system. Headnotes and key numbers have been singled out as determining factors in searchers' preference for the Westlaw system when compared with other information providers. The abstracts in the headnotes are customized to the legal domain, and the typical needs of legal users; they distill the text of the case to list of legal issues and descriptions of the discussions of these.

> S3: If [the abstract is] for a specific issue, if deals with a specific issue. Headnotes of document, I think for this one it would be important because you could keep west key headnotes and that's what I used to like about the whole west system that […] either you could search by that or you wouldn't have to read the whole thing, you could see all the things the court dealt with, all the different issues because even if the topic I was looking for wasn't the biggest or the most important issue addressed in that case, a lot of times it would have other issues that were addressed in the case. And then maybe mine would be at least, if the court dealt with it, it would be in that keynote, you could just read the keynotes without having to look through the case. And if mine was in one of those keynotes, it would be a lot easier to search from there just using that key number.

As the participant pointed out, abstracts are important time savers in information seeking, as they enable the searcher to quickly learn about the content of the document. This is very helpful while searching, and is even more advantageous when recalling a document record at a later time.

Further customization of abstracts can be achieved with user-created or user-annotated summaries. The topic of user notes is described later on in section 6.6. Notes to customize information would not only consider the special characteristics of the domain and the document (as headnotes do) but also those of the searcher, such as interests, tasks, previous knowledge etc. These abstracts are currently created by users in the form of notes and annotations.

> I: When you take notes when you are searching, how do you use those notes later on?
>
> P2: I usually use them when I'm writing the brief I'll refer to the notes, but I won't keep my notes forever.
>
> I: What's in those notes? Are they citations or keywords, or summaries?
>
> P2: Usually citations and summaries of the cases, but I think they might be better way of doing than abstracts.

Recording user-customized or user-created abstracts is very important because they represent how the results relate to the searchers' own problem and create a link to the task where the information will be used. These are usually recorded through user annotations and notes discussed in the next section.

6.4.3.4 Document text

Document text is a very likely candidate for reuse in legal information seeking, as discussed earlier. Saving the full text of the document found or making it retrievable to users is a very important function, as legal information is highly context-dependent. Even if only a section of the text will be used, the user may need to go back to the full text to check contextual factors. As discussed earlier, there should be a two-way link between the portion and the full text of the document. There should also be a link between the document text returned as a result, and the document written by the user where the result document is used.

6.4.4 Links among documents

Citation links between legal documents express many different relationships and are central to the content of the document. Two important relationships expressed through citation links are:

(1) court opinions, statues and other documents the current document builds on and cites

(2) subsequent history of the current court opinion, whether it was appealed, overturned, reaffirmed after it has been published. In the legal information domain these relationships between documents are especially meaningful and important.

When saving legal information, the links to and from documents should be saved as well. It should not require an extra step from users to save a hyperlink, it should be easy and included in the act of saving a case. In the following example the participant assumes that hyperlinks are an integral part of the document and should not be saved separately.

> S5: Hyperlinks, you can always get back to a hyperlink. So saving that information isn't… The whole point of a hyperlink is that you don't have to remember it. You are going to come back to it just naturally by doing your search again. So hyperlinks sort of save themselves.

Links are used in search techniques: in citation chaining and in citation analysis to look for documents that were cited by a group of other documents. In the following example the searcher uses links to find more materials related to a topic:

> I: Actually, the next thing I'm going to ask you is some of these things I will ask you why you would like to keep that and what you would do with that, so we can start with the hyperlink.
>
> P2: What I would do with that? Well, the links to the law review articles I think would be more if you were writing a trial brief or an appellate brief, they are usually written by students from law reviews, they are usually very good students. And it might give you some ideas about how to argue something. You wouldn't use it necessarily to cite it, but it might lead you to what's

important to narrow in the area.

Managing links from a document separately from the document is a need that arises from these search techniques, as the user may find a document that is exactly on point for his need and uses that as a starting point to find more documents by following the links from it. In this case the list of links should function as a checklist and help the user plan actions and monitor progress. Visited link color can serve that purpose in hypertext systems.

> I: when you print it out, would you use that in searching again?
>
> S3: yeah, I mean because what I may do, it's a lot easier to read it from here [paper] than from there [screen] and especially if you're being charged by how long you're on there, what I'd do is maybe read it, read the portion which is relevant and then if they cite other cases then I'd highlight those cases and then maybe print those cases out because a lot of time if it discusses a ... If this subject, this issue was very important to this decision, they would pretty much give a good discussion of the whole issue and that would give me a lot of the cases that I need to look at.

Managing links separately, highlighting, collecting, creating checklists out of links, as well as other data elements from documents are some of the functions needed in the user interface.

Recording what hyperlinks have been followed and showing this to the user can help with navigation, reducing disorientation

> I: Do you like that feature? It's all hyperlinked?
>
> S6: Yes, that's very good. Because the way my mind works in general, is I sort of travel like this stream when I'm doing things. And you go deeper and deeper, it is like creating a web. You go deeper and deeper looking for information and if you had to go reload, go back and forth rather than being able to keep following where you are intuitive sense is looking to go. Hyperlinks make that very helpful.

As hyperlinks can represent many different types of relationships, it is important to know what the relationship is. The Westlaw interface provided small icons next to items to show whether the relationship between cases is positive or negative. Whether a new court opinion has reaffirmed or overturned an older case is an important distinction for searchers. The type of link may influence what the document at the other end of the relationship will be useful for. The type of document at the other end of the link also influences how it can be used, a preview of that document before following the link can be helpful.

6.4.4.1 User-created links

Users can set up links between documents and sources, between tasks and documents, text sections and full text of documents, etc. These links can express type as well. Users path in the system can also be expressed and recorded by links.

> I: You take a lot of notes. If you could incorporate them with your search history, like here, can you put stickies on those?
>
> S5: I don't know, I never tried. No, I can't.
>
> I: […] How would you like to incorporate your notes with your search history?
>
> S5: By hyperlinks. I think that would work.

An interface tool to support assigning type values to links can be helpful for saving and managing search results and search histories.

6.4.5 Serendipitously found results

Serendipitously found results are unexpected benefits of browsing and searching. These are usually documents or information that may not be relevant for the current task but are important for another concern of the user. These need to be saved with proper indexing for the other topic in an easily accessible area in order to facilitate their use later on. If a document is relevant for another topic or task the searcher is working on, she should be able to save it for this other topic or task with a note on why it is relevant.

> Tyburski: I can't tell you how many times I have come across information on the Internet and said, "Oh, I am so glad to know that's there. I know I'm going to need that sometime." It may be the next day and I can't find it.

6.5 Search results management data

After locating relevant information to the searcher's tasks, this information is saved and prepared for use in solving problems. This process stretches across searching for and using information and involves many, often parallel, steps. This section describes information created while judging information found, and while preparing it for reuse. This process is described in detail in Chapter 5. It first, it involves a judgment of relevance of an item, then organizing the items according to the user's organizational scheme and annotating items with information the will help their use. As the focus of Chapter 6 is search history data, this sections describes relevance judgments and organized clusters of documents, while the next section concentrates on user notes.

The process of preparing information found for reuse is closely linked to user tasks. Relevance is judged in light of the searcher's tasks, results and often organized by tasks, and use notes relate the findings to the user's current knowledge and the problem that led to the information search. By allowing searchers to record this task-related information, the search history can support integration of the various subtasks of the larger information problem by creating a continuity between the steps that make up the larger task. In this respect, search and information use are treated as steps or subtasks of the larger problem that prompted the searcher to seek information.

6.5.1 Relevance judgments

Recording relevance judgments helps the searcher in using the information for his task. Often only those documents are selected and saved that are found to be relevant for some reason. This involves some kind of marking within the system and then a customized list or group of documents on the user's own computer. This is the simplest selection and recording of relevance. If a document is saved by the user it was found relevant, thus saving a document expresses a binary acceptance decision:

> S4: Acceptance/rejection?
>
> I: Whether you liked it or not.
>
> S4: That's irrelevant. What do you mean liked it or not?
>
> I: So you found an article and …
>
> S4: Oh, I'm only going to keep it if I liked it, so it's very important.

In reality a relevance judgment is not simply a binary yes-or-no decision. Some cases and documents are more relevant than others, some speak to one aspect of the topic, while others speak to a different aspect, or even a different topic or task. Some cases are good for the attorney's cause, while other are against it, but should still be collected and studied. Some cases are relevant, because they consider the same legal issue, while others are important because the defendant was involved in them. These different shades and colors of relevance can be recorded when the user thinks it is important to know. Attorneys generally described these decisions when searching and would sometimes record them in annotations or notes on the printouts or in the Word file. The following two participants described the importance of recording why the searcher thought something was relevant and to record whether the document saved for or against their argument:

> S2: A lot of times you can see, as a young associate or junior associate, you go to the partner and they'll say why did you pull this case? And you'll have to explain what caused you to think that this was relevant. And it's difficult to do that if you looked at a hundred cases without a road map.
>
> P2: And acceptance relevance judgments, we would want to look at all of the cases, even ones that might go against you, just to be able to distinguish them.
>
> I: So you would want somehow note that, that this is positive, this is negative, but I'm still interested.
>
> P2: Yes, because if you want to look at the.., usually the other side is going write its own brief and you want to look at the cases they cited and tried to why they are not applicable so you want to look at everything, even the ones against you.

A short reminder of the type of relevance can also help when searchers forget why they saved or printed a particular case. When entering relevance judgments, a pop-up window with a relevance template can help users in relating the information found to a problem or task. The relevance decision information will be used when applying the information found to a task, e.g. writing a document. This information may include pointers to how the information is supposed to get used. By showing why a case or document was relevant at the time of the decision the user's task of interpreting and integrating documents at a later time is made easier by reminding them of the previous interpretation.

Organizing documents into clusters around tasks can also help with reminding the user why a certain document was saved. By clustering it to a topic or task, it can give an indication as to what the searchers thoughts were when he saved the document. Clusters of documents are described in more detail below, in section 6.5.2.

Delayed relevance is when a searcher views a documents but does not think that it is relevant at the time of viewing it. The relevance judgment may change over time, the searcher can later reconsider his decision in the light of subsequent findings and the user may want to come back to the previously reviewed document. The change in the relevance judgment can be due to many factors, for example, later results turned out to be disappointing or information learned at a later point changed the relevance of the previous documents. In this case the searcher needs to be able to retrieve the items previously considered irrelevant. The opposite of this is weeding saved search results later. This happened in the following situation:

> S5: Apparently those laws had a lot of requirements to them and now this law is intending to simplify things somehow. I'm just looking at, maybe, by going to those earlier documents like I did, I can get some opinions as to whether this was a good law or bad law, because usually you get those opinions right when it becomes enacted.

When making a relevance decision and saving results, these can be linked to the task they will be used for right away. One case can be relevant for many different topics or tasks. Users should be able to note that on the case. Lin's work proposes a system to do customized indexing of documents found. His tool provides functionality to assign documents found to many interests or tasks of the user. Clustering document according to various attributes is discussed in section 6.4.2.

> I: So preserving what you were thinking when you were looking at that thing?

> S5: Yes. And then there was something else one page back that I thought was important. [client ID] It was … just because you may need to associate, later on at a later time, you may need to associate this information with that file and just because there so many thousands of files, you are going to need to … . You don't want to find really good information and then sit wrecking your brain trying to remember what this was pertinent to. If it was a particular file, just because offices these days just have so many files and so many clients. You might as well make a direct association right from the start, maintain that. Of if it to pertains to ten different files, you would need to keep all ten

numbers.

One of the attorneys in the study interpreted the recording of the relevance decisions as a rating system of how relevant a certain document was. He also suggested that the rating of documents would influence how long they are saved:

> S6: Acceptance/rejection relevance judgment. The rating system?
>
> I: Yes.
>
> S6: I'd like that in the current thing and I would also like that in the archives. And in the archives I have to say that I would be less likely to keep things that weren't a five-star rating. So I don't know if that takes this down a little bit, the value on the last level.

Relevance judgments change over time. As described at the beginning of this section, sometime users only realize later that a document would be useful. On the other hand, a document that looked relevant in the beginning of the search and was saved may turn out to be irrelevant by the end of the session, or the other way around. In the long run, cases may lose their relevance, and if storage is a problem they should be removed. Subject 2 describes his practice or marking changes in relevance decisions and explaining the reasons for them:

> S2: I generally keep .. if I print a case, it means that I really think it's relevant. If it turns out that it's not, I have a pile for thought relevant-not relevant, this is why I thought it was relevant, this is why it's not relevant. Then I have a pile of printed out relevant and highlighted, which areas were key. And then two separate piles with the search history connected to both.

The search history can record changing relevance judgments by allowing multiple relevance judgment records, each automatically dated and timed.

Weak relevance is defined as relevance judgments made about documents but not fully committing to the relevance of the document. If the user sees something that might be relevant, and marks it to come back to it later for a more in-depth investigation. Searchers need guidance to come back to these documents without forgetting about them. Marking these documents serves as a buffer to store them until later. One participant remembered the cases she wanted to look at:

> S1: All right, so I'm going back to the latter part of the search where I saw a bunch of cases that were from my court and I'll just pull them up quickly and I'll see if there's anything there. I tried to start where I think I'm going to get the most immediate gratification and if I find something in that particular court that would be very helpful.

The case, project or task identification need to be part of the relevance judgment record.

Some of the participants mentioned that before saving a document, they Shepardize it, as a way of making sure it is good law. This step comes before fully committing themselves to the usefulness of the document. Thus initial interpretation, relevance judgment and Shepardizing are typical pre-processing steps that are carried out before saving a document. The interface can prompt users to Shepardize cases before saving them.

6.5.2 Customized clusters of documents

Customized clusters of documents are groups of documents organized by some attribute, such as topic, author, or source. These clusters can emerge at various stages of the search, including pre-search notes, result gathering while searching, search strategy development, and information use. Saving documents attached to clusters proved too be a very popular idea among searchers. Clustering documents is a way of describing the personalized database of documents that a searcher builds up on their own computer from documents collected from various sources. Clustering documents around specific attributes can be achieved through indexing them by these attributes in a standardized fashion and then clustering them based on the index terms. Clustering occurs only for documents that were found to be interesting in the search for some reason. The participant in the following quote describes grouping result documents from a search around topics:

> P1: Clusters of documents would be really helpful, because I could have clustered all the documents I found that had to do with …, that one good Supreme Court case was not a textbook case, can't remember, oh it was the creationism statute, so I could have classed all the cases that talked about creationism but not textbooks. And then you have another one which talks about textbooks. And I think you are going to find those in different keynumbers. You are going to find them in different searches. So if you could group all those together, that would be nice. You could group them in a different way that Supreme Court case had to do with state statues, so you group them all by state statutes and you could group another by just local municipalities doing things.

Documents can be clustered based on a task as well. An attorney can collect all the documents related to a project or a brief to be written within a project and manage them together for the purposes of the task. Another example can be collecting documents to be saved together, in clusters. This is further discussed in the organized collection tool section, as a functionality to cluster documents within the result set as expressed by the participant in the quote above.

Clustering can represent subjective topic definitions of the searcher that do not match the subject categories or descriptions of the system, as described in the above interview excerpt where the searchers found documents in different searches and would like to assign them to one topical category.

Searchers should be able to assign documents to as many clusters as they wish, which is an important advantage of digital storage of search histories and results. This assignment can be automatically done based on common attributes among documents or can be defined by the user based on tasks or other subjective attributes.

> I: The clusters of documents, how would you… so you said you would want to keep them by subject matter?
>
> S1: subject matter.
>
> I: Is that the only clusters you would like to keep?
>
> S1: Probably, I might want sub-clusters within that. But generally it would be more by subject matter, I'm not. Well, let me back up. If I got a patent case, and I got a judge that I know only handled one or two patent cases, I might want every patent case by this judge, if this judge only had a few. Or if I'm working for a client that has been before this court or that judge multiple times, I might want to look at every single decision that involves my client, so I guess I have to sort of correct that, I might want sg broader than just subject matter of what case is about.

As the participant in the above quote describes, topical clustering may call for organization schemes that allow for multiple levels through hierarchical clusters.

The topical clustering structure may reflect the thinking of the attorney about an area of the law and later determine the outlines for future documents written by the attorney. The clustering structure should be easy to save and reuse for other purposes. In the next quote the participant describes why she assigned high importance to saving clusters of documents:

> I: The clusters of documents or groups of documents were also ranked high. When you find cases, would you want to group them into your own categories and keep them like that?
>
> P2: Yes, I think so, particularly if I were writing a brief, I would probably, I might want to break it down into how I was going to divide up the brief and which case is going into what subject area, so that would be helpful.

Incorporating the searcher's own documents in these clusters is also very important. This is described in Chapter 5, section 5.5.6.2 Building personal collections.

6.6 User notes, annotations and marks

Users took many notes, created frequent annotations, graphical marks, and highlights while searching in the study. Notes and annotations are text inputs. Notes are text records created more or less independently from information found, even though often inspired by it. Annotations are text records specifically linked to certain pieces of information found or the search history. The links between text records created by the suer and external information differentiate annotations from notes. Graphical marks have been used both in conjunction with search results or search histories, and independent of them. In the first case, they served as annotations or highlighter to draw attention to certain sections of the information or to present relationships between certain sections of the information. When graphical markings were used independently of search results and search history, they often served as an external representation of the searcher's knowledge model of a certain area, describing concepts and their relationships with each other.

Note-taking represented a step in addition to saving documents and text sections. Notes were explicitly created by the user, even though on occasion notes were created by copying sections of document texts. In this case, selecting and copying required value-adding actions from the user. This last method was used by the following participant:

> I: The first one is, what do you do that helps you remember your searches? And we did talk about your notes and the Word file.
>
> S5: Write things down, and mostly what I do is copy and paste into Word, just because it's just so easy to toggle back and forth. You just have to make sure you have Word open and start with a fresh document and just paste right into Word. Which is a lazy person's way of keeping track.

The format of notes can vary greatly. In the study, in some cases organizing materials with some graphical marking was described as notes taken to relay information to the searcher later on. Highlighting certain parts of histories in order to emphasize them and draw attention to them later was another valuable tool for searchers and need to be supported. Annotating printed or otherwise recorded search histories is another form of notes where the user attached handwritten notes to certain parts of the search history.

The initial processing of search results often involves marking, highlighting and annotating. As searchers read through documents, they relate them to their tasks and problems and interpret them. Their interpretations are then recorded in the form of notes and annotations, linked to the results. Thus the first level of interpretations and processing of new information requires marking and highlighting tools, and annotation functions. The participant in the next example describes this process:

> I: And then what would you with that document later?
>
> S5: I would save it and highlight it, or put a star next to the things that were particularly useful. Or I might write notes on the margin to go back to the whole document and not just to the segment I printed.

Participants often mentioned the need for marking, highlighting and annotating documents, systems and the search history for various tasks.

The annotations in this case link the results to the next step of the information-seeking process. They are often an intermediary step and sometimes destroyed after a new document is created, as described elsewhere.

Taking notes is different from annotations in that it is usually longer and not necessarily linked to specific places in the search history or search results. Notes can be taken to interpret information, as a preliminary format of a new document. Notes can also record planning information, this role of notes is discussed below in the section on task management. Notes can be taken through handwriting, typing, voice recognition or copying text from the search system or the search history window. When copying text, it is important for the system to "remember" the source of the text and create a bidirectional link between the source and the destination.

6.6.1 User notes

User notes represent the value added by users to the search history: it is not simply the record of actions, but the representation of the processing and value created by the searcher and thus search history amended by notes is more valuable and cannot be automatically recorded.

Planning, reminders, task management

Users bring a set of notes to the searching in most cases. These are notes they took before they started searching. These should be part of the search history as well, as they represent their thinking from the beginning of the information problem. One of the participants described his pre-search notes and they role in his searching as follows:

> I: Tell me about your notes that you started with.
>
> S6: These are from a conversation with a co-counsel who actually had some citations that were written down for us to check out. They were also some notes that I took on terms that I thought I was going to want to search. Keyword, as I was thinking about this how I frame some of the things that we were going to do.
>
> I: Is it basically pointers to what you …?
>
> S6: pointers and some directions and checklists to accomplish and use the time effectively, because as you know these things cost time and money. So you want it to be as thorough as possible and make your , see these types of searches.

Collecting and organizing results, results processing

Taking notes on results sets and individual documents, including graphical marks and highlighting, records searcher's relevance judgments, the circumstances of finding the document, and their interpretation of the information in the results.

Interpretation, recording knowledge

Recording results in a reformulated form can help the user with interpreting the information found and building it into his old knowledge structures. Writing notes that relate the new finding to old knowledge or the problem at hand also help with this.

One of the participants described this type of user notes as the following:

> S2: Notes are crucial, because most people can't remember all the thought that went into their projects. You do research so that you can give somebody an answer.
>
> I: So, that means that when you read something you want to reformulate it to make that answer?
>
> S2: Yes, you need to understand how that fits into what the question is and is that going to get me to an answer.

In this case it is very important that the notes are linked to the part of the search history they are interpreting.

6.6.1.1 Content of user notes

The task of the user influences what he records in his notes. The task determines what the notes will be used for, which in turn affects the nature of the notes. The content of notes vary greatly, notes can be made about the following objects, among others:

(1) documents

(2) sources

(3) queries

(4) search as a whole

(5) topics, issues

6.6.1.2 How long to keep notes

Length of storage of user notes depends on the type of note. Immediate planning notes can be destroyed at the end of the sessions. One of the participants reported describing her notes on new knowledge and interpretation, as she only used these to help her learning. Notes taken to support document writing can be destroyed when the document is created:

> I: when you take notes when you are searching, how do you use those notes later on?
>
> P2: I usually use them when I'm writing the brief I'll refer to the notes, but I won't keep my notes forever.

6.6.1.3 Context of notes

It is important to link the notes to when and where the user made them. Notes are inherently context-sensitive, they are thoughts inspired by some part of the search history, and the may contain information pertinent to another part of the search history. Linking to the step or information in the search history that the notes relate to helps interpreting them later:

> S6: Ok. Notes would be the same thing. I would like to be able to … would that be different than graphical marks?
>
> I: Yes. So that would be typed.
>
> S6: I would like that. I would it typed like a little popup window next to a certain case or something, or file, to put a little comment box, that would be great.

It is also important to know the circumstances among which they were created, the person who wrote them, etc. The study participant below describes attaching metadata to notes created in the word processing file, which creates a specific type of note, a note about a search:

> I: And what kind of thing do you put into that Word file?
>
> S5: Well, this one I just printed out, but what I might have done was gone to the top and typed in some thing, I might have typed in today's date, the client ID, whether this was an easy or difficult search engine to use. If I pulled up like a hundred, and I only looked at fifty, I may say there looked at 50 of 100. Looked at first 50 or something like that.

Notes created while searching are an integral part of search histories and should be recorded along with the automatically recorded search history information. These also need to be organized, as shown in the following example:

> I: So you create your little notes in Notepad application on the computer?
>
> P1: And then file them, very old-fashioned.

6.6.2 Annotate

Annotations are user notes that are linked to a specific document, query or other search history item, as described above. They are linked to a specific location in the system, the search history or the results. In addition to marking, annotation stood out as a crucial tool in customizing automatically saved search history information to user needs and making it more accessible for searchers. Annotation and highlighting ties search history information to the user's tasks and thinking and it prepares information for reuse, thus linking searching for and using information. The interface must provide a flexible tool for annotating objects at varying levels of granularity. Saving the annotations as linked to the locality where they were created crucial in reusing them.

Study participants often mentioned a preference for printed copies of documents citing ease of annotation and highlighting as a strong deciding factor. Paper is a known medium and tools for marking paper are widely used.

> S2: at least before there was no way to save your Westlaw searches. So, by printing out your search history, your terms, connectors, and everything, instead of trying to recreate the wheel, you would know. And it was a lot easier; especially I used to highlight my search history, if you can… When you go back and you want to update you can actually see like the dates of the cases, and all that stuff. So it's a good … it's a road map.

Annotated printouts are easy to share. While an automatically recorded search history may be difficult to interpret, if it is marked and annotated by the original searcher, it can be more easily shared. Tools for annotating and sharing search histories and results online would reduce the need for printing and paper-based collaboration.

> S2: I used to search to see if certain insurance companies had filed certain products, and you know we know that this company is coming out with something and it's new, and I'd go and I'd do a search everyday to see, and I would print the results, highlight the name, wherever that company name showed up. And I would say not what we're looking for. So that would let the partner know, that you not only did the search but you looked to make sure that it was not

Organizing search histories and search results can be closely related to annotation. Sometimes searchers use annotations to express organizational structures, and when they assign an item to a category they often take notes on why the item fits into the category and what its relationship is to other items in the category.

Annotations on results

Searchers can also record the result itself in a shorter, more customized form, if the system does not provide adequate results representations, as in the following quote:

> I: And now I just have a couple more questions, when you do research, how do you currently record your search history? Do you take notes about your searches?
>
> P2: I take notes, which is not best way of doing it, as I see. And if I have an actual trial, I'll usually Xerox the case, I'll take the book and Xerox the case.
>
> I: Ok, Xerox the case that's relevant to your case?
>
> P2: Right, yes.

In this case the system should support moving some information from the results to user notes and complementing that with graphical marks or annotations. Annotations often contain relevance judgments, discussed in the relevance judgment section below.

6.6.3 Mark

The majority of the attorneys described using graphical marking in searching and information processing. In the following examples Subject 3 and 5 described what they would do with printouts of the text of a case after finding it and deeming it relevant:

> I: Once you print a case, what do you do with it?
>
> S3: Generally read it, make sure that it is a good case, and then highlight it, you know, use it, I'll cite it, in the notion, or whatever I'm writing either verbatim or I'll just use it as a general cite. It's good to have a copy for the future if I need to use the case in the future, if it's relevant.
>
> I: And then what would you with that document later?
>
> S5: I would save it and highlight it, or put a star next to the things that were particularly useful. Or I might write notes on the margin to go back to the whole document and not just to the segment I printed.

Users put marks documents and search histories and search histories in order to emphasize or draw attention to certain parts for future use. The information marked will very likely be reused or revisited in some fashion, the goal of the marking is to make retrieval or return easier.

> S1: […] typically, when I get sidetracked like that, I'm on case eight out of 15. And so I'll Shepardize eight, eight will be helpful, I'll Shepardize eight, I'll find another case that's helpful, I'll Shepardize sg there, I'll find another case that's helpful, and then I want to go back to nine out of 15. Once I've sort of explored that tree to its fullest extent, I want to go back to number nine, cause usually at that point I haven't read it, sometimes I would write it down, eight is very helpful and put a little star by it and go to no. nine to 15, see if there is anything of interest there and then go back, […]

In other cases the graphical marks help interpretation by highlighting important ideas and expressing relationships between units of information, such as circling and linking text sections in a document. These types of marking also appear in user notes, as described by Spink et al.

Highlighting and marking may also serve to designate certain parts of a document or system as a function. An example is to highlight a part of a document to mark it as the title or designate a part of the system as an opening location or as a location that will be recalled at a certain action.

Bookmarking is a special type of marking where the searcher marks a place she wants to come back to alter, this type of mark will also register in a list of marks to provide easy access for the user on her return. This type of marking is described above in detail, including the need to bookmark smaller-than-page units of digital information. Improvements on traditional bookmarking can be made in order to better fit the needs of users. Marking sections as relevant to a certain topic or task can be one useful function, as described in the section on relevance judgment decisions. In this case, the marking can not only represent a link to another item, but can also show the type of the relationship, in case of relevance judgments: the reason the document is relevant, is it for or against the attorney’s case, the extent of helpfulness of the case etc.

6.7 Length of storage of search history items

The length of storage is influenced by legal and ethical responsibilities of legal practitioners as described above. The type of information also influences it. Timely information that changes often and needs to be updated is less likely to be saved, but the queries and sources to access this information are more important to record. Data and results that do not change very often can be saved and reused if the need arises again.

Ethical concerns may influence the length of storage as well, as described by participants. Records are required to be preserved for a certain amount of time by the state and often the organization itself. Keeping records after the required time may be a problem.

As storage space is becoming more and more affordable, keeping records for a longer time period should not be a concern with the expense of storage. Thus, keeping search history information in the visible, surface layers of the system, accessible to the user. The older the history gets, the more it can get aggregated, with more detail provided for recent history. The sections of the search history that are “hidden” can still be accessed, but viewing them requires extra actions from the user.

The type of information system may influence the length of storage as well. The web as a medium changes frequently even if the content does not require updating. This was even more important in earlier stages of web development when website and pages were less stable. A frequently changing medium would require the saving of content in case the source would disappear. Pointers, queries, URLs and other reusable information is also important, but these may change.

6.8 Granularity of saved search history and results

Information saved from searches varies along the granularity dimension, from keywords to document sets in the search results domain and from steps or pointers to multiple sessions when saving search histories. The user should be able to save large clusters of documents just as easily as attribute sets or selected sections of full text.

> I: When you search nowadays, and not just legal but whenever you search, what do you do to help yourself remember things. Do you create reminders of any kind?
>
> P1: When I’m searching? Just generally?
>
> I: Yes.
>
> P1: I write myself, I create a document in Notepad and write in there.
>
> I: What is it that you write down?

> P1: Either web addresses, or I just copy part of the document that I was interested in and put that in Notepad, something like that. Or more how I did search, is that what you wanted to know?
>
> I: I'm more interested in what is it that you would record about your search and how you would record it, so anything.
>
> P1: If it was on the web and I found one link that got me to all these other things that were really helpful, I'd record that one link and then I'd write down "Got me to blah-blah" But other times you don't really care about the links, you just care about what the information is. So I'd just write myself a little note that says this is in there.

The systems should allow saving various granularities of information and also deleting sections of the records in varying granularities. In the following example the searcher just wanted to save a section of the document:

> S5: I think it's printing the whole document.
>
> I: Oh, you just wanted that page?
>
> S5: Yes. Oh well, I'll just grab the whole document.
>
> I: Next time we can select the part of the text and then just print that that selection.

Options to save results at varying granularity would help searchers with collecting better focused results. The user should be able to save large clusters of documents just as easily as attribute sets or selected sections of full text. In search history displays however, it serves the user better to bring all objects to the same granularity for easier comparisons.

7 Findings: Interface design

7.1 Introduction

The theoretical framework and the empirical findings on the role of search histories form the basis for designing search history interfaces. Providing a continuously growing history record in the user interface is the most direct use of search histories. Interface design recommendations for displaying search history data are presented in order to feed the recorded information back to the user. In addition to direct search history displays, tools building on search history data can also help searchers in search-related tasks. Search-history based user interface functions are also described organized around a scratchpad and a results collection tool.

Section 7.1 recapitulates the main search history use areas are introduced (Section 7.1.1), then briefly outlines interface implications (Section 7.1.2), and reviews the Object-Action Interface (OAI) model (Section 7.1.3), as this model is used in the rest of the chapter to describe interfaces.

Section 7.2 describes the eight physical interfaces developed and critiqued during the Phase 2 participatory design sessions organized around the process that was used in Phase 2 and present participant reactions to them. The eight interfaces are divided into two groups:

(1) Three interfaces (VictorWEB search history, MdUSA search history, Internet Explorer history) were used to introduce the search history concept to participants. Although these were not meant to be critiqued, participants expressed their opinions about them; these critiques are summarized along with brief descriptions of the interfaces in Section 7.2.1.

(2) Section 7.2.2 discusses five interface designs, partly external designs, partly designed by the researcher. They are presented here in the order they were discussed in the participant sessions. They were presented to the participants to be critiqued, the descriptions of the interfaces and the participant reactions are discussed in three separate subsections: 7.2.2.1 - 7.2.2.5.

Section 7.3 describes conceptual interface design guidelines. While section 7.2 summarized the process of designing and critiquing interfaces in Phase 2, Section 7.3 presents the lessons learned through the process in 7.2. Some of the information is repeated here, but is organized around conceptual guidelines, as opposed to physical interfaces reviewed and critiqued by participants in section 7.2. Section 7.3 presents interface design guidelines and recommendations for search history displays and two search history-based tools: scratchpad and organized results collection tool.

Section 7.4 discusses two general history-based interface guidelines not linked specifically to any of the three conceptual interface design tools described in section 7.3. When describing interface features in section 7.2, participant comments are often clearly distinguished. In 7.3 however, where conceptual interface designs are presented, Phase 1 and Phase 2 data are presented uniformly with the researcher's insights and the source is not distinguished. This is done in order to increase the readability of the text.

7.1.1 Main search history use areas

Previous chapters described what user tasks search histories can support (Chapter 5), what information to record (Chapter 6), and how to manage this information (Chapter 6). The three main user task areas where search history information can be helpful are:

(1) Information-seeking task management: This area is comprised of user tasks related to the planning, execution and evaluation of information-seeking actions, including search strategy development and task integration from searching to information use and back. It also includes finding and collecting information and finally recording action information for the purposes of billing, keeping statistics, and justifying decisions is also included here. See sections 5.2, 5.3, 5.4, parts of 5.5 and 5.7.

(2) Knowledge integration and management: While the first task area factors are focused on many physical actions, this second area describes the many cognitive steps that make up the individual searcher's knowledge management and mental model development while searching for information. Many of the interface tools will overlap between the first two areas. See sections 5.5 and 5.6.

(3) Collaboration: Sharing history information helps asynchronous collaboration by creating a shared action space, even though participants did not physically or temporally co-locate. See Section 5.8.

These areas can be further broken down into specific functions provided in the user interface. The following section describes three interface solutions, all of which support these three use areas.

7.1.2 Interface implications

From the exploration and definition of the higher- and lower-level use tasks, the following three interface tools emerged as important:

(1) Tools for presenting multiple views of the search history and for search history representation; tools for searching for specific elements in the search history. When annotations and markings incorporated: search maps/trip logs, showing history of items, comparing items, repetition of items

(2) Scratchpad functions with annotation, note-taking and highlighting capabilities, next to the search screen where searchers can copy and

record things and take notes. Note-taking, annotations, highlights and graphical marks: Customizing search results, task integration into info use, interpretation.

(3) Tools for organized results collection that will be the first step in using the information #1 (e.g. writing a brief)

In addition to these tools, generic history-related design guidelines inform user interface design.

When considering the application of search history information, it is very important to keep the users' work tasks in focus and not to overburden users with too much history information. Information will help the searcher achieve their goal, but is not the main purpose – the attorney is focused on writing a brief or defending a client.

7.1.3 Object-action interface model

Several methods of modeling user tasks and actions and user interactions with interfaces are described in the human-computer interaction literature. Card et al. (1983) describe the GOMS (Goals, Operators, Methods, and Selection rules), which is a cognitive engineering approach to identifying and describing user tasks and needs. In order to describe user interface design needs in terms of search history information, Shneiderman's (1998) Object-Action Model was selected as it closely matched the structure of the research questions and findings. The research questions in Chapter 3 divided the research into three large areas:

(1) What user tasks can search history information support and how? (actions)

(2) What search history information to record? (objects)

(3) What user interface tools and functions are need to allow the user to use the search history information in support of his tasks? (objects and actions)

Shneiderman (1998) described the syntactic-semantic model of human behavior in order to describe programming, database management, and direct manipulation. Semantic knowledge of users describes meaningfully-acquired semantic concepts such as the act of saving documents in word processing. Semantic knowledge is meaningful in the context of the task domain and was later described as task-domain concepts. Syntactic knowledge concerns low-level actions, rote-memorized syntactic details such as how to delete a line or what steps are needed in order to save a document in a specific application on a given computer platform. Semantic concepts refer to tasks well-organized and stable in memory while syntactic details can be arbitrary and required frequent rehearsal.

This distinction has led to the development of the object-actions interface model (OAI) over time, especially with the advance of direct manipulation techniques, which also led to the minimization of syntax in user interfaces. The OAI model consists of task objects and actions and interface objects and actions. An examination and understanding of task objects and actions enables the designer to create metaphors to represent these in the interface. The objects and actions are organized into hierarchies. An action, such as saving a document can be broken down into finding the "File" menu, clicking on "Save" once, entering a file name and clicking on the "Save" button once.

Graphical user interfaces and direct manipulation led to the minimization of syntax in user interfaces. In many systems, users need to remember device- and system-dependent methods of executing actions – in the word processing scenario they need to remember which series of actions saves a file and which action deletes a line. Learning and retaining this knowledge is difficult for users, as it is system-dependent and varies widely across systems. Minimizing the need for this type of knowledge should be a goal of interface designers which can be achieved through direct manipulation methods and the application of standard widgets and familiar metaphoric representations for task objects and actions.

7.2 Interface design development and critique process

The OAI model is used in the following sections to describe the various interface designs demonstrated to participants or designed by them. The interfaces include screens from existing systems to familiarize participants with the search history concept, interfaces designed by the researcher, and other search history interfaces to be critiqued by participants.

7.2.1 Basic search history interfaces (Introduction to search history concept)

In the participatory design session, participants were first familiarized with the search history concept through a discussion and brief demonstration of three search history tools:

(1) The University of Maryland Libraries' VictorWEB Web-based online public access catalog search history display. (Figure 7)

(2) The University of Maryland Libraries' MdUSA database access interface search history display. (Figure 8)

(3) The Internet Explorer (IE) history, version 5. (Figure 9)

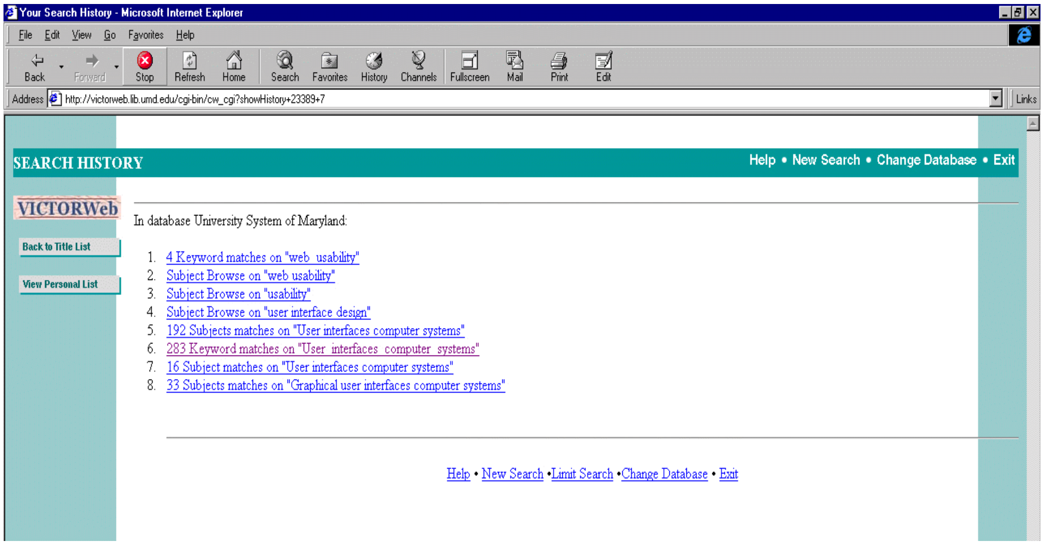


Figure 7. University of Maryland Libraries VictorWEB search history.

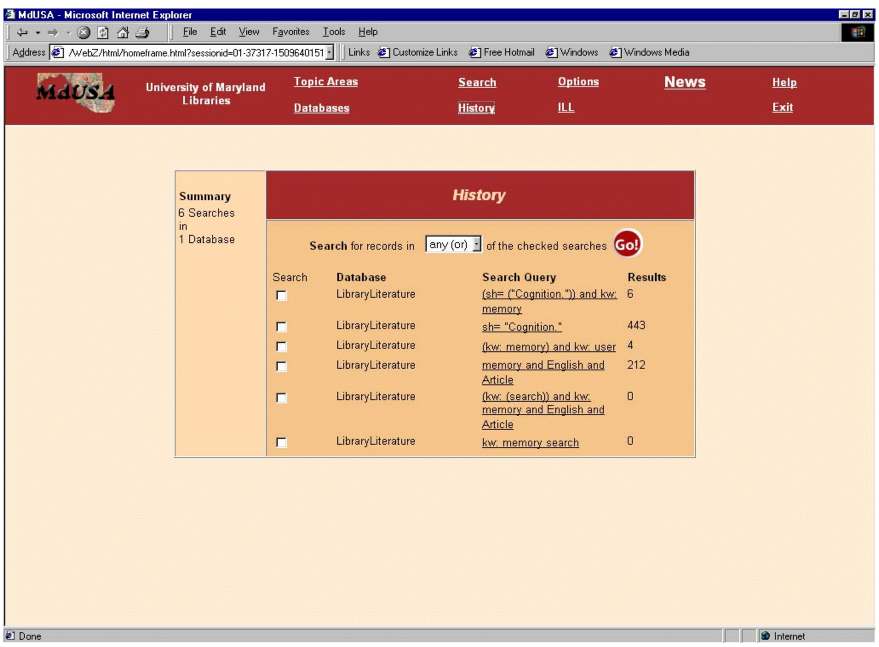


Figure 8. University of Maryland Libraries MdUSA database access interface search history.

Figure 9. Internet Explorer history.

As stated earlier, these interfaces were presented to introduce the search history concept. Even though participants were not specifically asked to critique these designs, they expressed their opinions. Participants remarked that they did not understand the relationship between the lines in the VictorWEB and MdUSA histories, where the queries were represented by those text lines. The VictorWEB history presents information about searches, their type, the query terms, and the number of results returned, but the information given is limited. The MdUSA history displays the name of the database as well, since in this system the searchers could search in multiple databases. In addition, it allows users to combine searches from the history list and rerun them in combinations. Participants saw this as an important feature; they wanted to be able to rerun queries from the search history list. In viewing these interfaces, one participant remarked that another system did not record which database the query was run in and this caused a major inconvenience for searchers when the queries were rerun in a default database. Another concern expressed was the length of time search histories are stored for. The IE history was generally popular, especially the function that lets users search in the history. Discussions of bookmarks highlighted problems with organizing and finding bookmarks and the ability to access the individual's bookmark list from any computer.

7.2.2 Advanced search history interfaces and history-based tools demonstrated to and critiqued by participants

In Phase 1, several interface design ideas emerged. These were realized using low-technology prototyping tools, such as paper, HTML and Visual Basic, and presented to experts. Integrating the experts' suggestions, new prototypes were created; these, along with other search history interfaces, were presented to participants in Phase 2 for review. The following five interfaces were presented to the participants with instructions to critique them.

Text interfaces:

(1) Text design 1: Westlaw Search History Trails (Westlaw designs, 7.2.2.1)

(2) Text design 2: Westlaw Search History Trails Add-on Tools (Dissertation designs, 7.2.2.2)

Graphical interfaces:

(3) Graphical design 1: Ariadne time line of search histories (Twidale et al. designs, 7.2.2.3)

(4) Graphical design 2: Simple time line (Dissertation designs, 7.2.2.4)

(5) Graphical design 3: Complex time line with results (Dissertation designs, 7.2.2.5)

The interfaces were grouped by text versus graphical designs, and within both groups they were arranged in the order of increasing complexity. The text designs displayed textual lists in temporal order of history items: events and documents. The graphical designs used the same temporal order, but presented and arranged history items visually.

Participants were not told which screens were designed by the researcher in order to encourage them to express their criticisms. Eventually, participants were also asked to "fix" or redesign these interfaces or create their own solutions for the use of search history information. A more detailed discussion of the methodology used can be found in Chapter 4. The results from these various processes are presented here through a review of the interfaces and participant comments.

The results are described in terms of medium-level interface object and action relationships. Search history task-domain objects and actions were described in the chapters on search history use and search history data.

7.2.2.1 Text design 1: Westlaw Search History Trails (Westlaw designs)

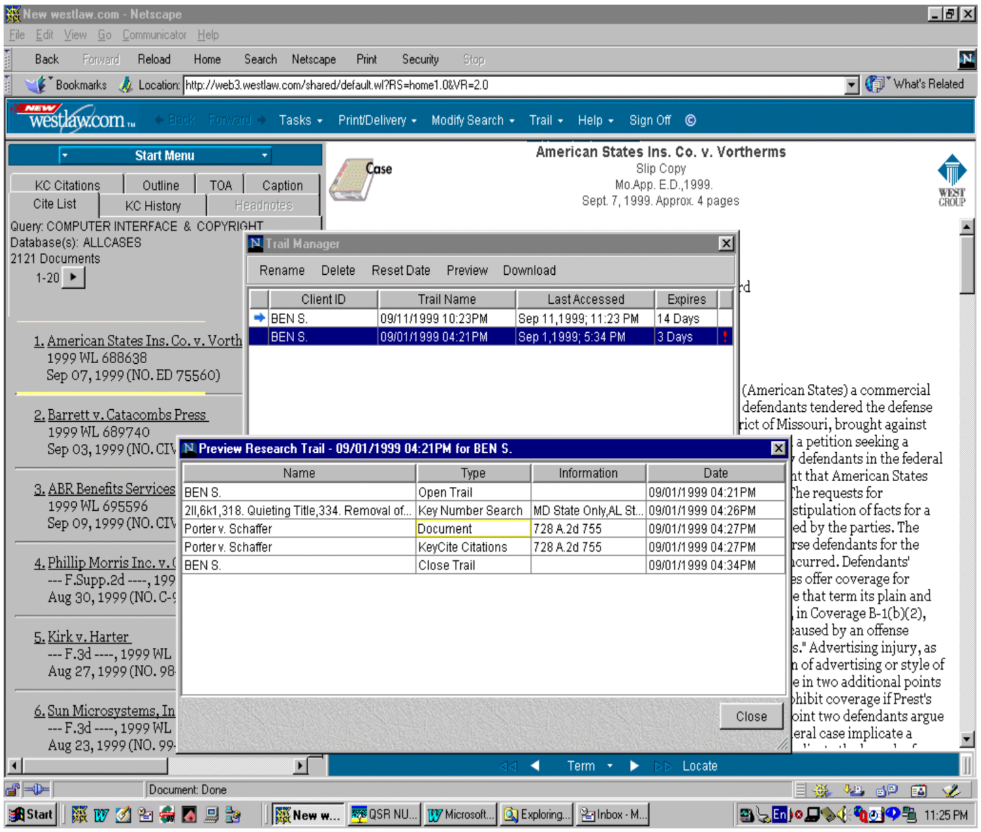


Figure 6-0. The Westlaw.com search history interface

Figure 10. Westlaw Trails search history.

The Web-based Westlaw system (Figure 10) at the time of Phase 2 data collection (September-November 2000) provided the text-based Trails search history feature in the search interface. As shown in Figure 10, the system captured most user steps and presented these in a text list. The objects and actions in this interface are presented in Table 9.

Objects:	Actions:
(1) Trail (made up of:)	(1) View trail
(2) Research event (terms): query terms, case titles, key number and title, etc.	(2) Start new trail
(3) Information (some type of metadata): database where query executed, case citation, database where key number search executed, etc.	(3) Open old trail
(4) Research type (type of action): search, find, Keycite history, etc.	(4) View old trail
(5) Date: date and time of action	(5) Download trail
	(6) Trail manager: Rename
	(7) Trail manager: Delete
	(8) Trail manager: Reset date

Table 9. Objects and actions in the Westlaw Trails search history.

In addition to the automatically created trails, searchers could name their trails and open and close old trails. The system stored the trails for 14 days, after that searchers had to reset the date of the trail or save it in HTML format.

Participants generally liked the availability of this tool, although they were confused by the sequence of actions. They remarked that the column headings were too generic and did not present the information clearly. Several participants mentioned that the information provided was not clear on how the user got from one step to the next; this proved to be the most important problem with the design. For example, when a case title is displayed with a Keycite History Research Type, it is not clear whether the case resulted from a Keycite History action or Keycite History was the action taken on the case. In the Westlaw system the second is meant by the line. As a general rule, it was suggested that in case of every unit of information both the action that led to it and the action that was taken on it should be presented.

Clicking on a search history line took users back to the original screen where the action was taken. This feature was highly regarded by participants. The lines representing queries took the user back to the results screen which was a problem for some of the participants. In general the opportunity to get back to both the query and the result list seemed optimal for most participants.

Another critique of this interface focused on the display of ancillary paths, such as a Keycite history action on a case that was returned to a query. The Keycite history action was interpreted as a tangent to the main task of queries and should be displayed at a lower level, indented or in a different color to express the branching from the main task line.

Participants remarked that the system did not display the number of results next to queries, which is a very important piece of information. The Westlaw system did not display this information, nor did it present result cases that were not looked at. In addition, participants remarked that cases should display where they came from, the fourth case out of 25 should from the query on "copyright" should present this information. Both of these pieces of information help the user understand the process better and also remind him that there were more cases in the set to be reviewed. The Westlaw trails did not record printing and saving actions, which participants would have liked to see.

In reaction to the trail management functions, such as saving, opening and closing trails, participants expressed a need for easier switching between active and passive trails, and the ability to have multiple trails opens and easy switching between these as this represents real-life task scenarios better, when users have to work on multiple tasks in parallel.

Another need mentioned was the integration of multiple databases, systems and sources in the history interface.

7.2.2.2 Text design 2: Westlaw Search History Trails Add-on Tools (Dissertation designs)

Interfaces for organized results collection tools were developed as an add-on to the Westlaw Trails interface based on the data collection in Phase 1. First paper mockups were designed and tested, then a Visual Basic interface was prepared. The three tabs represented three large functionality areas that emerged from the attorney observations and interviews:

(1) Gather search results and search history data

(2) Compare search results and search history data

(3) Search search results and search history data

These tools appear on the right side of the Westlaw trails interface and allow searchers to drag and drop items between the two windows.

As these prototypes were created as an add-on to the Westlaw Trail tool, they encompass all the objects and actions from that the Westlaw Trail system with additional objects and actions. The additional objects and actions in all three tools (gather, compare, search) are summarized in Table 10. Objects and actions that were added on to the Westlaw Trails interface are in bold.

Gather

The gather tool (Figure 11) is an electronic equivalent of the paper folders and filing systems attorneys currently use, with major extensions of functionality to take advantage of automatic logging and electronic documents. Searchers can set up and manage folders, save documents, parts of document, notes, searches and other parts of search histories and results in the folders.

This physical Gather interface design was used to learn about this functionality. The lessons learned here are later presented in the conceptual interface design sections 7.3 and 7.4.

Objects:	Actions:
(1) Trail (made up of:)	(1) View trail
(2) Research event (terms): query terms, case titles, key number and title, etc.	(2) Start new trail
(3) Information (some type of metadata): database where query executed, case citation, database where key number search executed, etc.	(3) Open old trail
(4) Research type (type of action): search, find, keycite history, etc.	(4) View old trail
(5) Date: date and time of action	(5) Download trail
(6) Folders, folder name	(6) Trail manager: Rename
(7) Document full text	(7) Trail manager: Delete
(8) Document section	(8) Trail manager: Reset date
(9) User notes, annotations	(9) Create folder
(10) Document sets	(10) Move documents, search actions (from search history to folders, between folders)
	(11) Merge folders
	(12) Divide/split folders
	(13) Order items within folder
	(14) Order folders
	(15) Color/highlight folders or items
	(16) Annotate
	(17) Compare
	(18) Search

Table 10. Objects and actions in Gather, Compare, and Search tools. Items new as compared to the Westlaw Trails search history are bolded.

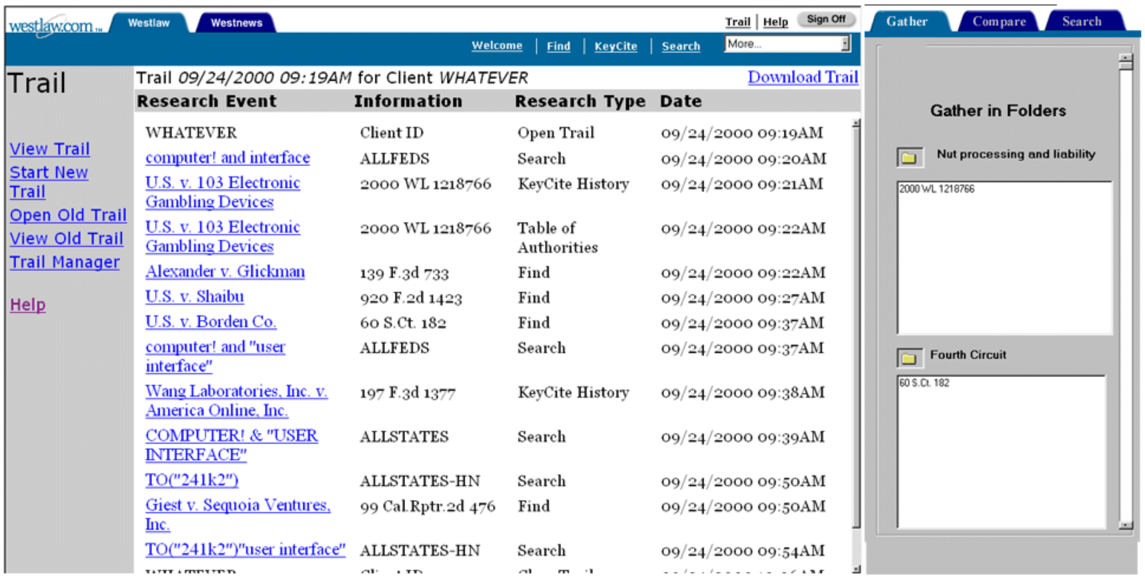


Figure 11. Gather tool as an add-on to the Westlaw Trails search history.

Participants suggested that search history actions such as queries and Keycite history actions could be moved to folders in addition to search results. These search actions can also serve as starting points for folders. The search results may be organized around the queries; users should be able to name folders by assigning parts of search histories to them. Functions for organizing items within the folder and organizing the folders themselves should be provided along with the organizational tools.

When transferring results of a query to a folder, the system should give the user a choice whether at the time of retrieval she would like to see the result set as it is or whether the query should be rerun at the time of retrieval and the most up-to-date result list should be presented. An update function related to each query and other action is another approach to address this problem.

As described earlier, attorneys often want to organize research folders according to multiple attributes: client, topic, chronological order, level of court and so on. This is more easily done in the computerized environment than in paper-based files; users should be able to duplicate items between folders with retaining relationships among the duplicate copies. When saving results items, a pop-up window can help them with indexing the item. The window can provide the main attributes of items in the searcher's personal collection, and allow the user to enter or mark these attributes. The attributes can include task and topic as main organizational avenues, where the searcher can attach the items to categories.

Flexible folder management tools are necessary to enable users too change folders as their interests, knowledge of the area and amount of materials collected changes. Creating new folders, merging and dividing folders are functions participants were especially interested in.

Changes to folders and folder content can be automatically recorded as part of an expanded search history; this was mentioned as an important feature by participants. An automatically added date attribute can help searchers track changes and versions. This is an example of task integration.

The search history folders should seamlessly integrate with the filing system of the user, it should save the search history and the results into folders that are also accessible through the file explorer interface. It should allow the user to select the place in the filing system where the folders should be saved.

Annotations should be allowed on the folders through text notes that are attached to certain folders or items in the folder or sets of folders or items.

Compare

The goal of the compare tool (Figures 12 & 13) is to allow searchers to filter out overlaps and differences between various result sets. Moving results sets to the similarity or difference areas should result in a list of overlapping cases between the two result sets. This tool allows searchers to manipulate the result sets returned to their queries as a search technique. Although more complex Boolean queries can achieve the same goals, end users often require a more direct management of the results sets than going back and reformatting the queries. This tool was less enthusiastically received by participants, they did not see the direct value of the ability of comparing, although this may be attributed to the lack of detail on the prototype.

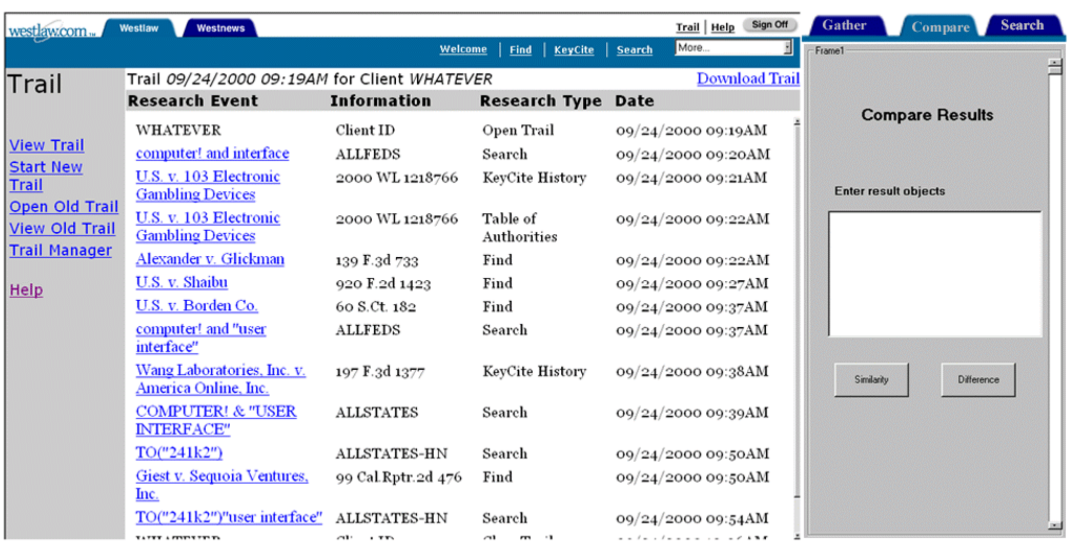


Figure 12. Compare tool as an add-on to the Westlaw Trails search history. The text box represents a filter where user can enter items to compare.

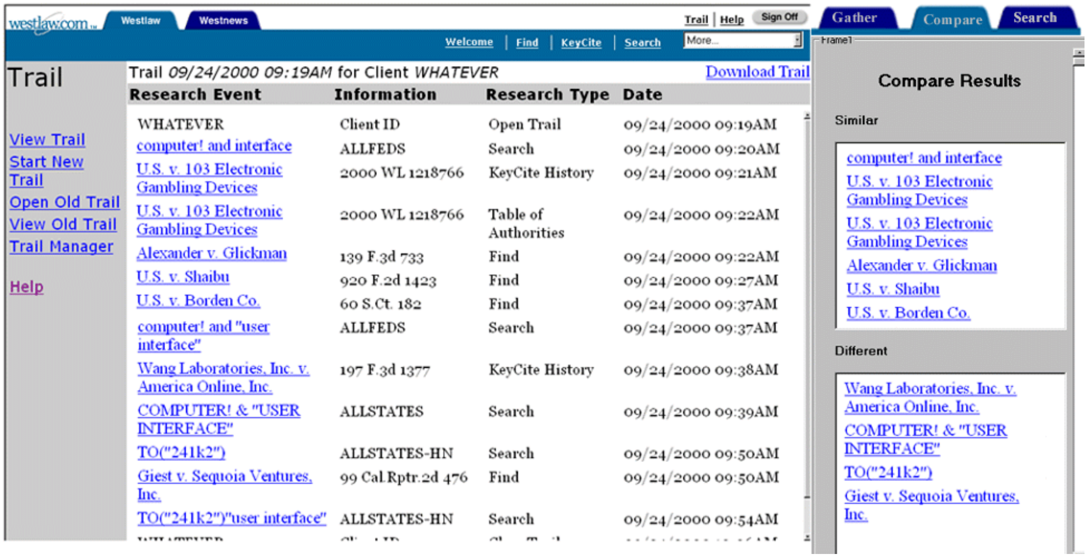


Figure 13. Compare tool as an add-on to the Westlaw Trails search history. The results of the comparison of two document sets are presented in two lists of similar and different documents.

Search

The search tool (Figures 14 & 15) allows users to search their search history and saved results records. This tool was also available in the IE interface demonstrated to participants in Phase 1. Participants liked this tool; they wanted almost all attributes of search histories and search results to be searchable. Search intermediary participants emphasized the need for an one-line search box for this purpose where searchers can enter complex Boolean queries. Presenting the results by user-defined clusters was also mentioned by users as a feature.

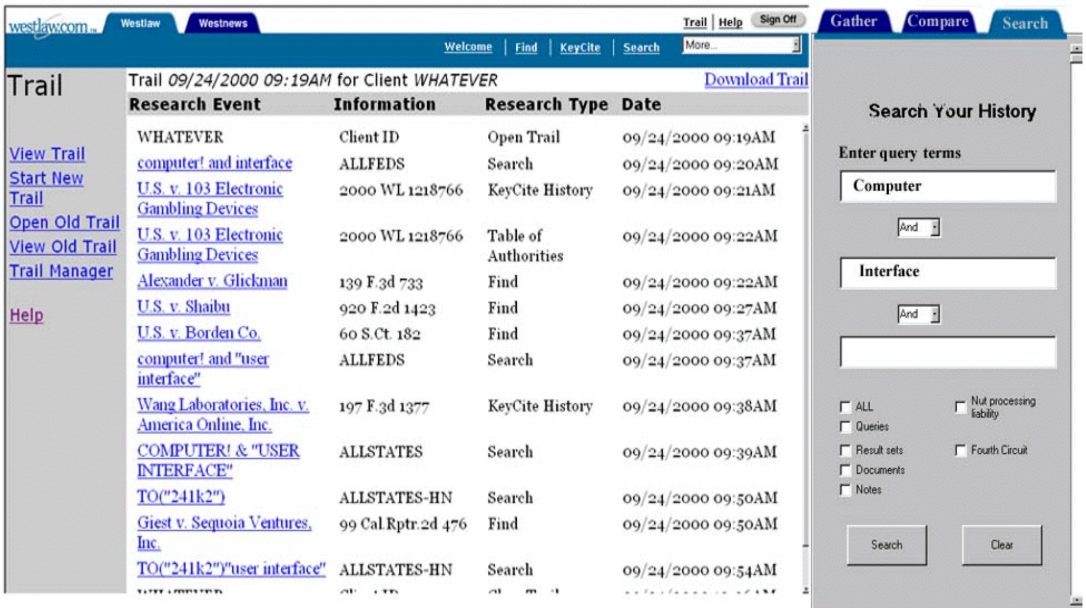


Figure 14. Search entry screen of the search tool, an add-on to the Westlaw Trails search history. The user is searching the search history for the terms: "Computer" and "Interface".

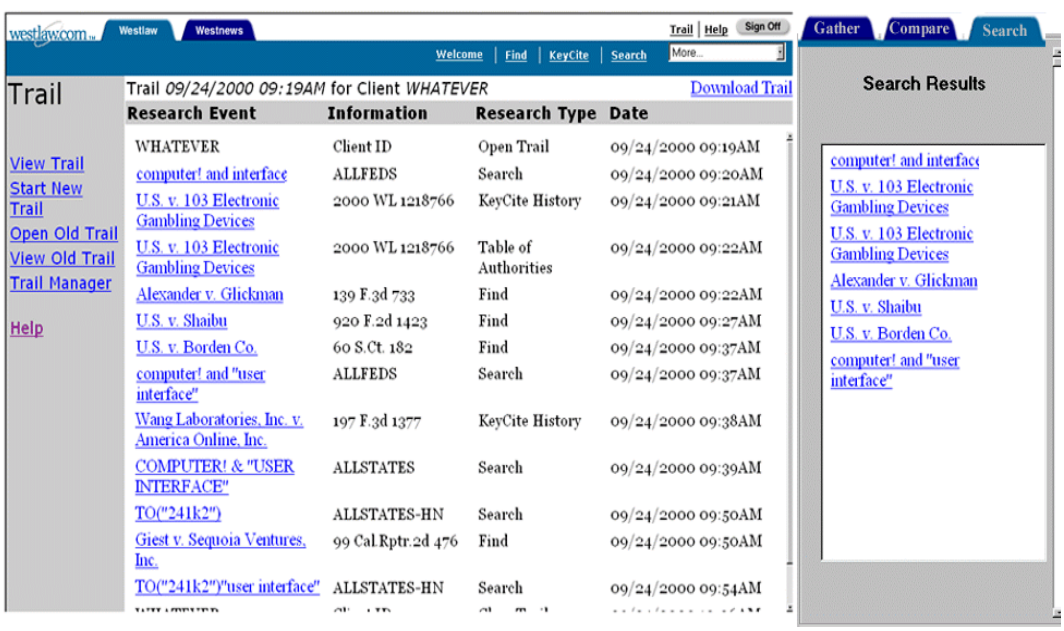


Figure 15. Search results screen of the search tool, an add-on to the Westlaw Trails search history.

7.2.2.3 Graphical design 1: Ariadne time line of search histories (Twidale et al. designs)

The Ariadne interface (Figure 16), described earlier in the literature review, was presented to the participants as a canned demo. This was the transitionary step from the text designs to the graphical designs. It gave participants an example of a graphics-based interface, as opposed to the Westlaw interface which is text based. Three functions were presented: pop-up screen shots, annotations, and the fold feature. Preference for graphical vs. textual search history displays seemed to be based on individual characteristics; various groups of participants expressed strong feeling for and against graphical displays. (This aspect was not specifically tested.)

Both types of displays can show structure. **Structured displays were always favored**; even proponents of text-based displays would have liked to see (1) indented lines to represent tangential actions and (2) structures and information to represent the relationships between steps. Most participants described interfaces that were both textual and graphical: textual information expressed the specifics of actions and documents, displayed query terms, showed document titles, while the graphical information presented this text information in an organized manner expressing relationships and structure.

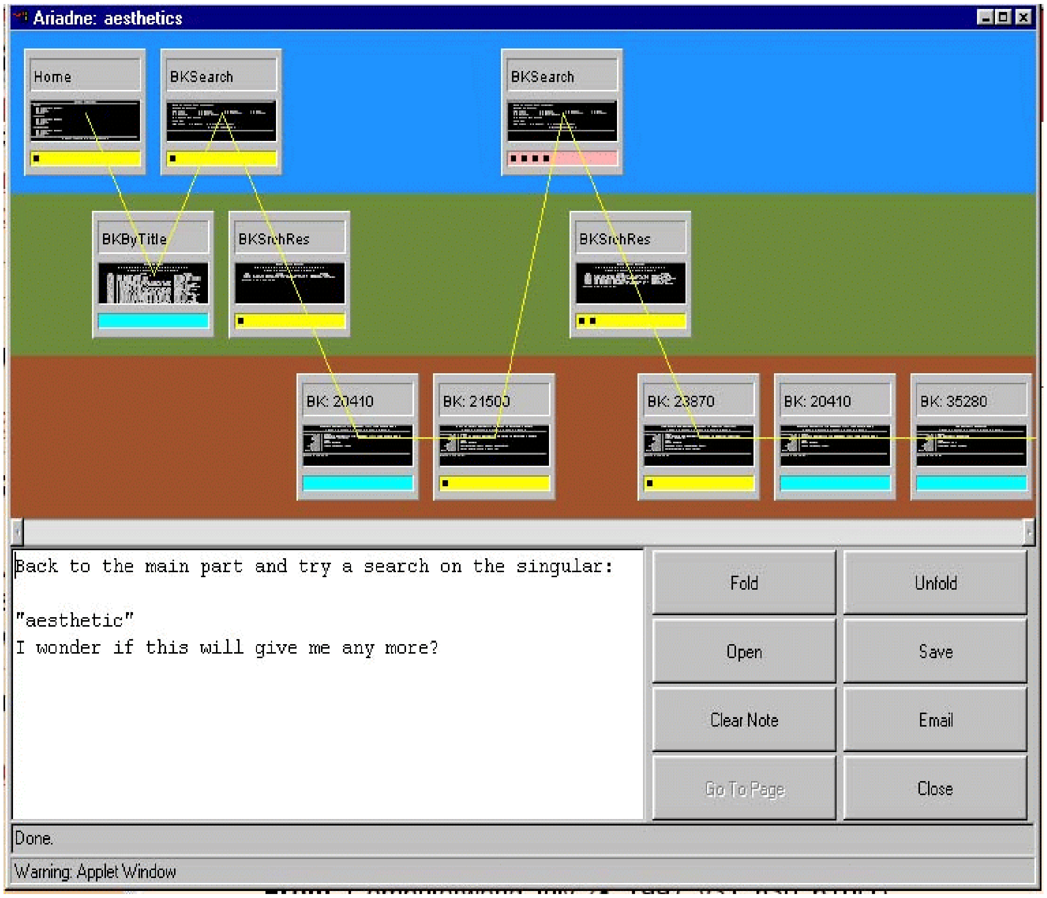


Figure 16. Ariadne search history display. (Twidale and Nichols, 1996)

Objects:	**Actions:**
(1) Queries	(1) Fold
(2) Search result set	(2) Unfold
(3) Documents	(3) Open
(4) Annotations	(4) Save
	(5) Clear note
	(6) Email
	(7) Close

Table 11. Objects and actions in the Ariadne search history interface.

Table 11 presents the objects and actions in the Ariadne interface. Combinations of these objects and actions make up the functionality of the interface.

Information presented

Participants commented on the lack of result set information. As with previous designs, participants required the display of the number of results returned to a query, along with the number of documents viewed from the result set. Presenting these numbers helps the users understand unexamined tangents. Another concern mentioned earlier and reinforced here was the lack of specific information displayed. Although searchers could tell at a glance how many queries and documents were on the screen, they had to click extra to find out what the document title was. They disliked this solution and wanted high-level representations of the actual content, for example keywords for the queries, so that they can at once tell what the queries and documents were.

The ability to traverse back to an original document or action in the history list through clicking on their history representation was mentioned as a requirement.

Annotations and marks

The annotation feature was widely successful among participants, although the colored lines representing the existence of an annotation and the dots representing the length of the annotation were again deemed too general; keywords from the annotations would have been more informative. The ability to take notes and attach them to actions or documents is a very important feature according to participants. In addition to text annotations, graphical marks to highlight important documents, steps or paths are also needed.

The fold feature enables users to select to distinct steps and “fold” the steps between them: make a section of the history hidden. While the technique was confusing for participants, they generally expressed a need to be able to hide parts of the history and thus aggregate the actions.

Scale up - graphic display question

Participants expressed concerns about the graphic nature of this history display: they were worried about performance issues with large graphics displays on computers, and were also questioning the ability of the design to scale up to long search sessions.

7.2.2.4 Graphical design 2: Simple time line (Dissertation designs)

Another time line of history events (Figures 17 & 18) applied to legal searching was developed based on the data collected in Phase 1. After paper prototyping and expert review, a prototype using a computer graphics package was created and demonstrated to participants.

Two levels of the linear search history display are described here. They are presented in order of complexity, power, and amount of information displayed. The first interface tool presents the events temporally. This display occupies a small section of the screen, it can be placed at the top or the bottom of the screen. User annotations and markings are facilitated by sticky notes that can be applied to any part of the search history and by highlights. Functions are described in Table 12 in terms of interface objects and actions.

Objects:	**Actions:**
(1) Database	(1) Search
(2) Query	(2) Preview
(3) Document	(3) Zoom
(4) Key number	(4) Annotate
(5) Keycite	(5) Aggregate
(6) Annotations	(6) Collect
(7) Document cluster	(7) Move
(8) Shopping/collection basket	

Table 12. Objects and actions in Time-ordered search history display.

Its main purpose is to provide :

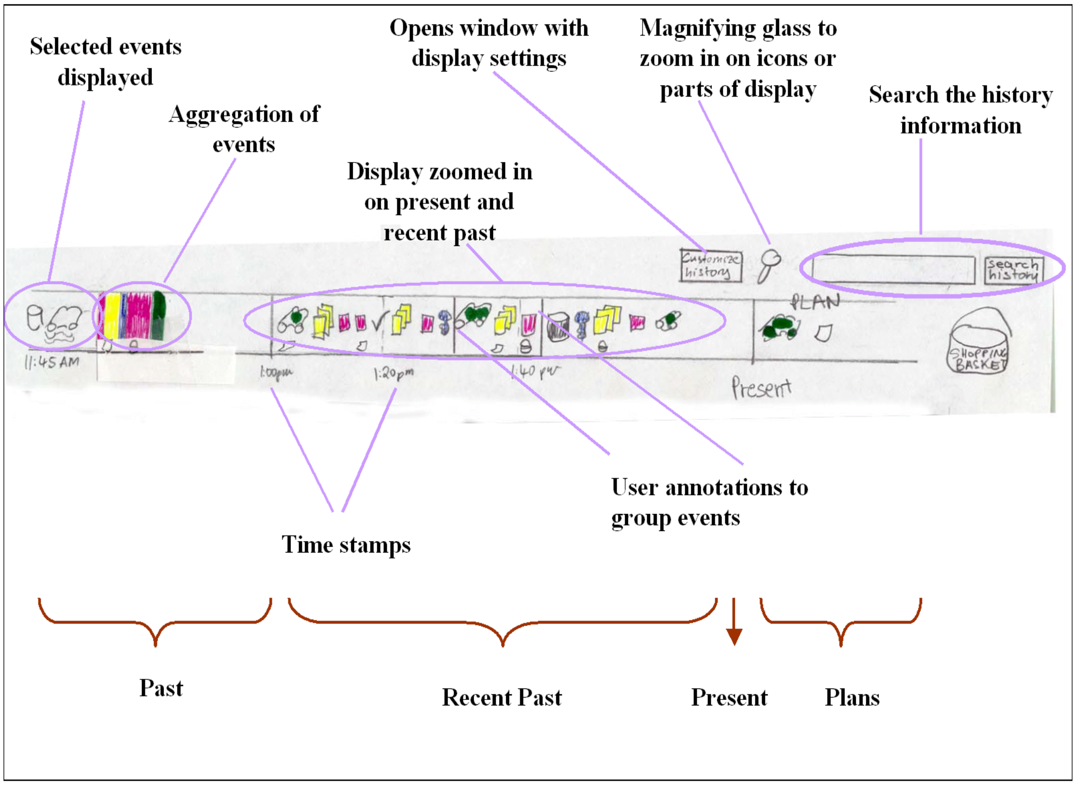
(1) navigation aid to go back and forth between actions,

(2) undo and redo possibilities,

(3) working memory aid with easy access to previously displayed information,

(4) simplified visual overview of the process of searching

(5) direct access to previous system states.



Level 1 interface: Timeline of information-seeking events.

Figure 17. Time-ordered search history display.

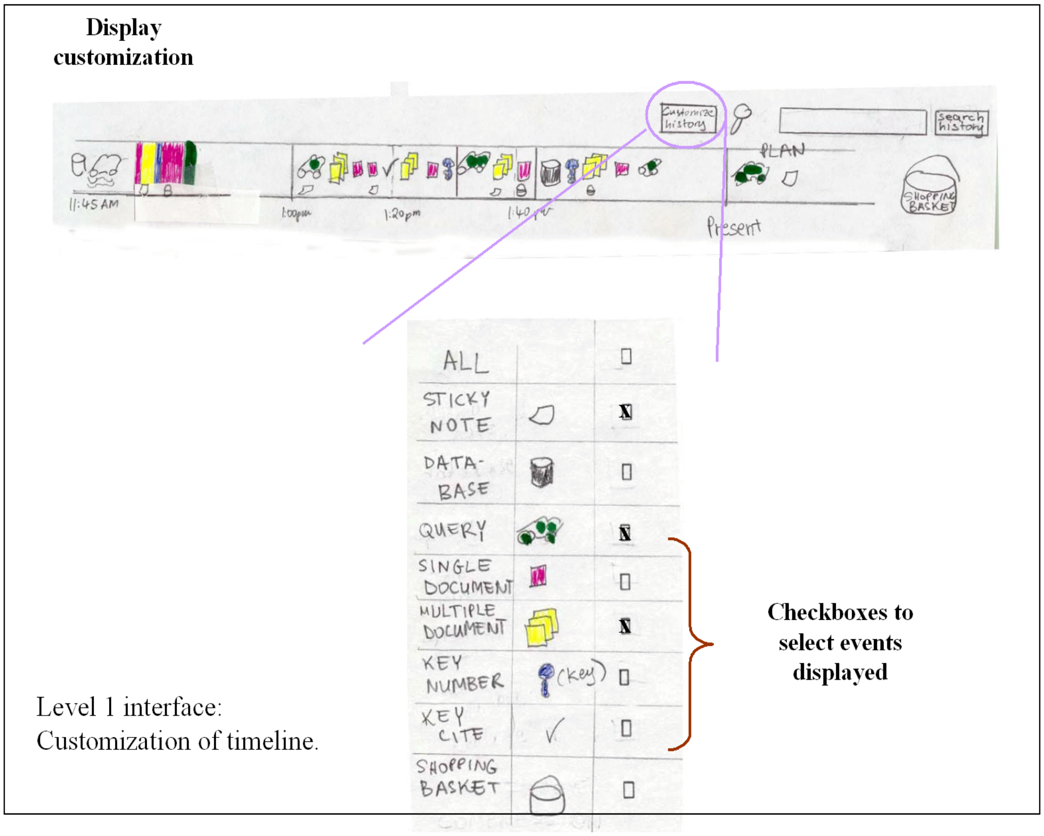


Figure 18. Time-ordered search history display with legend.

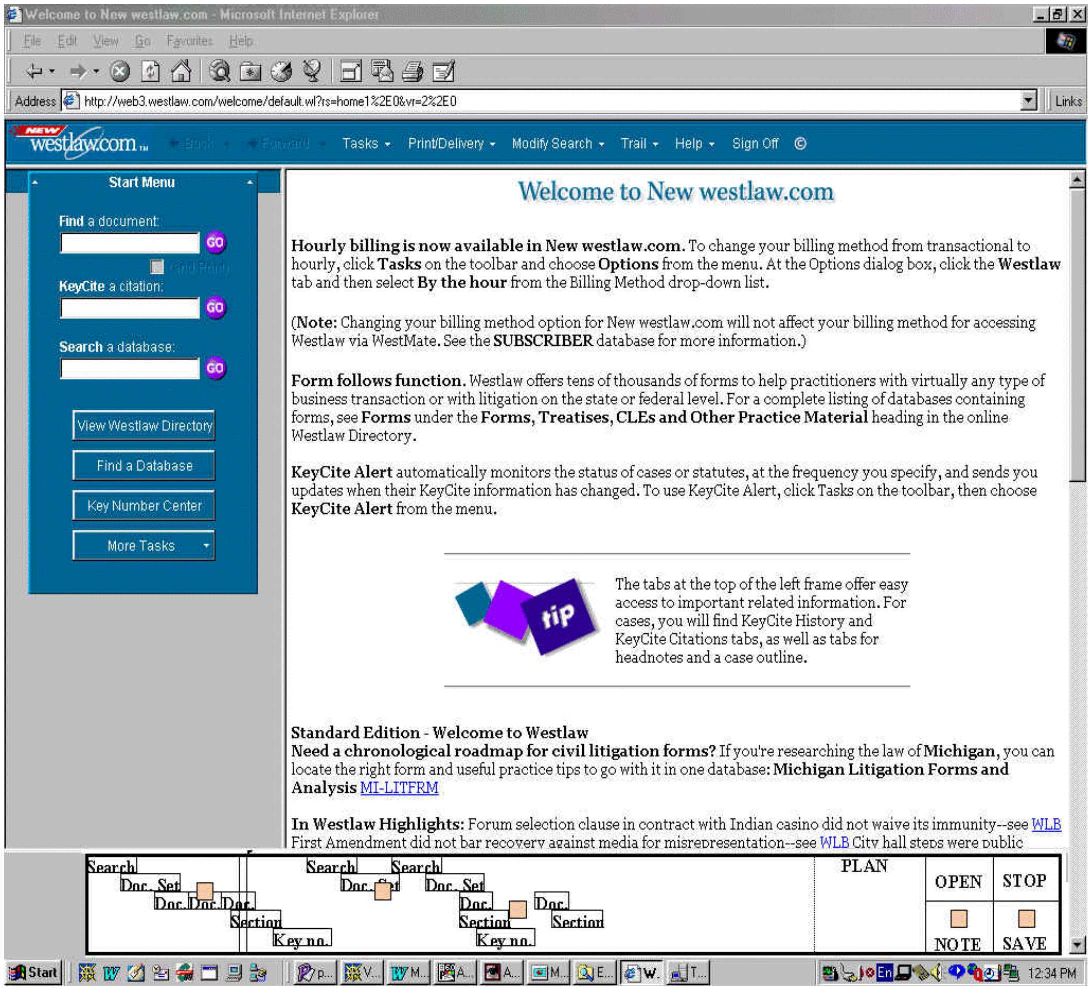


Figure 19. Time-ordered search history display integrated with Westlaw screen.

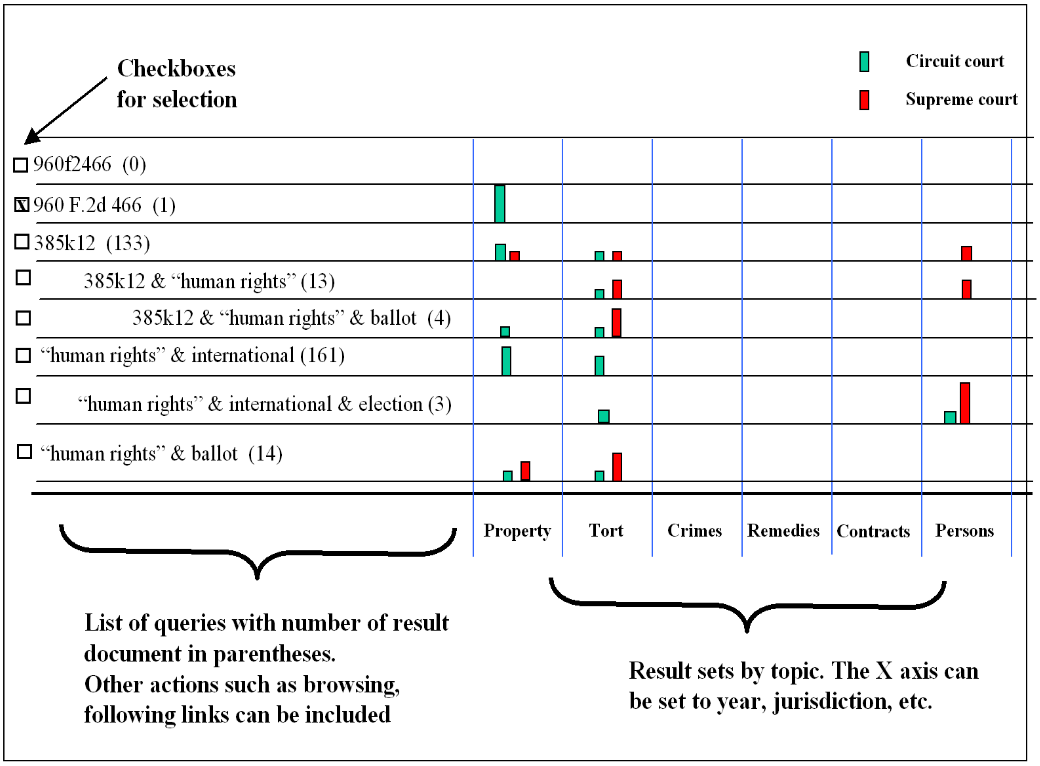
Figure 19 presents the version of the simple timeline used in participatory design and critique sessions. This interface is based on a time line with fish-eye, magnifying glass, or other zooming mechanism for the user to focus on selected areas. The default view is focused on the present and recent past of the time line, earlier past is contracted to save space but still provide an overview to this part of the record. Every action the user takes is represented by an icon on the time line. For explanation of icons see Figure 17. As the time represented and the amount of history information grows, the earliest events are aggregated into higher level representations or are filtered down to a specific types of action (e.g. the sequence is partitioned into searches and only the beginning point of searches are represented, but the user can view the details on demand). Once the information cannot fit on the screen even in aggregated form, the time line becomes scrollable. Scrolling over the event icons will result in a small pop-up window with some information about the event. The user can zoom in on contracted sections of the history and see more details, move events back into the search screen repeating that step. She can stop and restart recording, pause the recording to explore tangents. Users can search the history information by entering keywords into the text box on top of the time line. They can drag the magnifying glass tool over icons to get more details than in the mouse-over pop-ups.

Relationships between events have to be represented to show the branching of the action tree. Visual features of the icons can represent this, for example if a query is a modification of a previous one or if a document was examined before slight variations of the event icons can be used. One danger here is to represent too many different things with the same icons. Figure 18 shows the prototype that was presented to the users, where some of the features described were illustrated.

Participants remarked that the information presented about the items was too limited and general: specific search terms and document titles represent the steps better. In addition, they would like to have easy previewing techniques. Although they liked the at-a-glance overview of the events and the easy access to previous system states by clicking on steps in the history, they thought that finding out more about the steps this way takes too long and attention and two different parts of the screen. When designing the interface, rollover popup windows were suggested for the preview option and described to the participants. Some participants even required two levels of previews: a brief and a detailed option for each item. Some of the participants suggested a solution where the history items would be displayed both on a graphical time line and a text list that were tightly linked and selecting an item on one display would highlight it on the other, combining the Westlaw trail system and the graphical time line.

Comments on the graphical display included the need to collapse and expand the hierarchy to focus on selected sections of the history. One participant remarked that while she would not use the tool during searching, she would like to use it for post-search results analysis and summarizing results for patrons in an intermediated search environment. Providing flexible viewing options with search history displays is very important as user goals determine the type of display needed.

7.2.2.5 Graphical design 3: Complex time line with results (Dissertation designs)



Level 2 interface: Integration of search results.

Figure 20. Time-ordered search history display integrated with Westlaw screen.

Objects:	**Actions:**
Query	Expand
Set of documents	Collapse
Topic	Preview
	View

Table 13. Objects and actions in the Time-ordered search history and results display.

Figure 20 represents a more complex solution to presenting search history information in a different temporally-based display, where the relationship between actions is better represented. Here queries have been selected as the beginning points of actions, and all other actions are related to queries. The actions of a user are grouped and displayed hierarchically to facilitate browsing different stages of a search and getting an overview of the actions. Queries are indented under the first query if they are reformulations of the previous query - broader or narrower.

This window can be displayed along with other search and browsing windows and automatically updated as the search progresses. In this interface user actions are closely linked to their results. On the left side of the screen a hierarchical tree structure, similar to the one in the previous example, shows the semantic relationship of actions. The user has the choice whether only the hierarchical action tree is displayed or the results as well. They can serve different purposes at different stages of information-seeking.

Relationships between actions are richer than simple tree structures. As mentioned before, a document can appear in multiple result sets, queries are often modified and other relationships exist between elements of the history information. Interface tools are needed to show these. The functions are analyzed through objects and actions in Table 13.

Participants generally thought this display to be overly complex and inappropriate for an attorney's tasks which are usually tied to strict time restrictions. They saw this tool as more appropriate for a detailed analysis of results for a Law Review journal article, or academic work. The topical layout proved to be the most important problem, participants were confused by how the topics were selected and what they stood for. This illustrates the general principle that different history displays are needed for different purposes, user groups, and tasks.

Another display desideratum is to provide temporal visualization of search results in a topical hierarchy. In this part of the display searchers can select different queries and represent their result sets in relation to each other. A timeline of queries can be included in the display with sliders where the user can dynamically view results as he sets the timeline, or select multiple queries and show those at the same time. Presenting results of multiple queries in the same topical hierarchy give users two dimensions of information-seeking criteria. First, it is defined in the query using traditional information retrieval techniques, such as full-text Boolean searches with all its flexibility. Second, when visualizing the results of queries by dynamically comparing result sets, users can relate their original query to the topical hierarchy that represents the conceptual map of the area. Figure 20 presents a legal topical hierarchy with links to items returned in response to a query. This display can be developed further to show multiple queries along a time line.

7.2.2.6 Participant Designs from Phase 2 Participatory Design Sessions

Some Phase 2 participants were asked to design their own interfaces, below are several examples of the designs created.

In the first design (Figure 21), the participant presents search results organized by source, as she explains, she remembers searches by the source. The tabs at the top of the screen list the various information systems she uses to find information. Bookmarking often-used sources was mentioned as an important improvement to the search system display. When a tab is selected, a list of available databases in the system is displayed. Search history data is organized into query-based units, after the database is selected, the list of searches run in the database appears in a collapsable hierarchy. Each search is represented by the date, client matter number, search terms and the number of search results; the list can be reordered by any of these attributes. The frame on the right provides a search tool to search the history information. The participant anticipated using the information when similar queries are submitted.

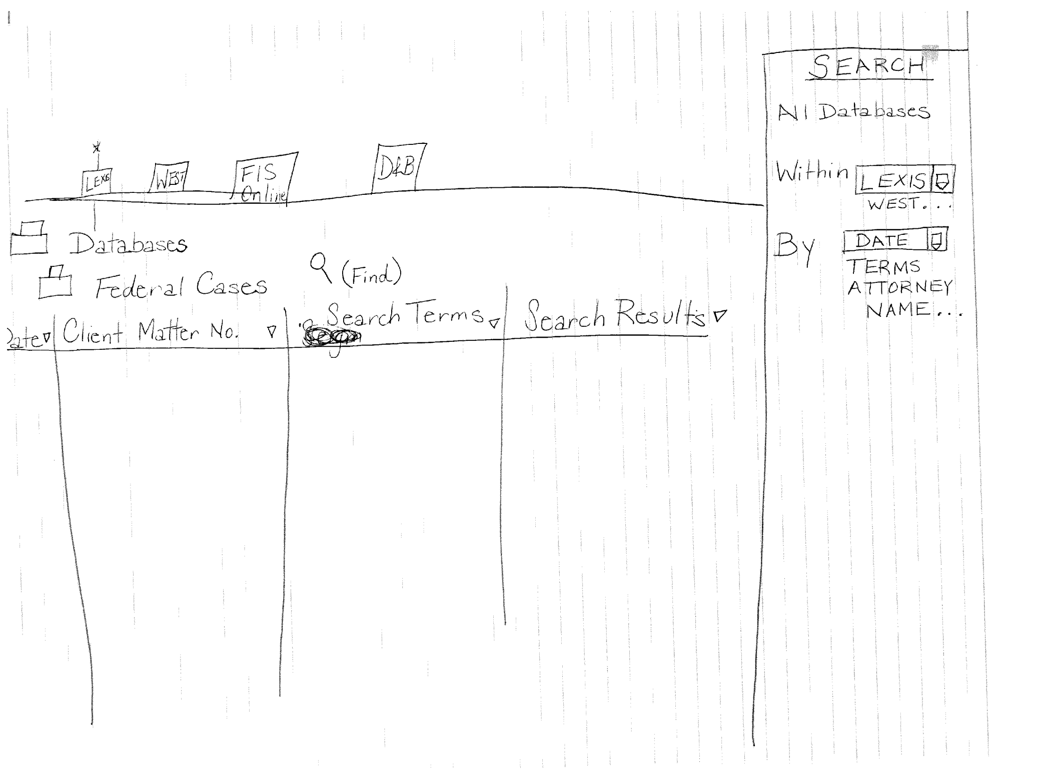


Figure 21. Participant design 1.

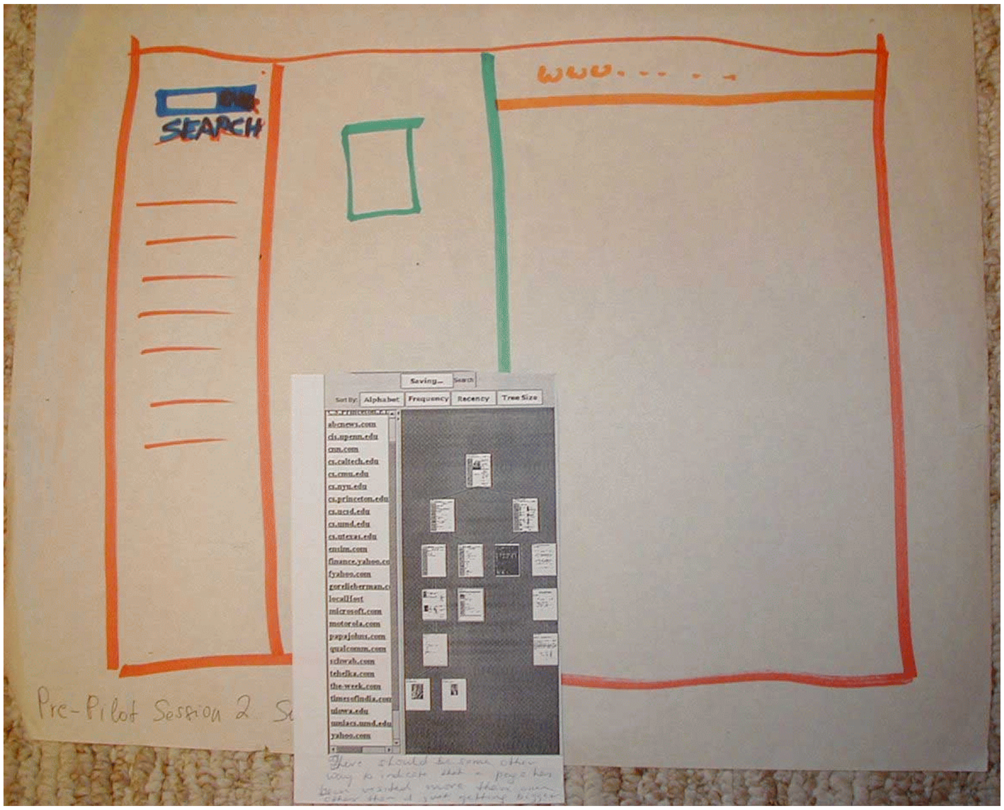


Figure 22. Participant design 2.

In the second design (Figure 22), the goal of combining textual and graphical displays can be seen. The frame on the far left contains a textual list of the history, while the middle frame presents the same history in a graphical view.

Both of these designs were created by participants introduced to the Victor, the MdUSA and the IE histories, but not the other history designs. Participants who were introduced to the more advanced history designs described combinations of those and did not create original designs.

7.3 Conceptual Interface Designs

Section 7.2 presented the process that was followed in Phase 2 data collection. It discussed the physical interface designs that were used in Phase 2 to elicit participant feedback. This feedback was included in Section 7.2 organized by the interfaces used. This section discusses the data that emerged from Phases One and Two in a unified fashion organized around three conceptual interface design phenomena:

7.3.1 Presenting search history displays

7.3.2 History-based scratchpad tool

7.3.3 History-based result gathering tool

Section 7.3.1 summarizes the wisdom gained through the process of presenting five search history interface designs in the process described in section 7.2. Sections 7.3.2 and 7.3.3 describe history-based interface tools, as opposed to direct search history displays. Data collection in Phases One and Two highlighted the need for tools based on search histories but not presenting direct search history displays.

7.3.1 Presenting search history displays

In describing direct search history displays, this section first discusses the history representations themselves (Section 7.3.1.1), and then comments on tools to manage these (Section 7.3.1.2).

7.3.1.1 Displays

In observing and interviewing legal searchers, and reviewing and designing search history displays with design participants, several important design principles surfaced and were discussed in detail in the search history use chapter and the first part of this chapter. These are summarized here and presented as guidelines for the design of displays presenting search histories to searchers while looking for and using information.

Search histories are composed of user actions, objects such as documents and other types of information, and relationships among these. A user action is often called a step in this dissertation. A step is made up of an action and an object.

Generic design guidelines derived from the data collected are presented in Table 14.

Design Principles:

(1) Align search history with user tasks.

(1.1) Different views of history needed for different purposes.

(1.2) Parallel search histories for parallel tasks: keep multiple search histories/trails open at once, with easy switching between them.

(2) Screen real estate, previews, overviews, and zooming.

(2.1) Do not make the history display rule the search system screen. The search system is a tool to support searching, not a tool to manipulate history.

(2.2) Provide easily accessible previews of history items.

(2.3) Provide overviews of action sets and sessions, based on chunking of steps along with the ability to zoom in on parts of the history.

(2.4) History item icons should link back to the event they represent. Make functions with specific content available.

(3) Type of data shown

(3.1) Types of data to save were described in Chapter 6. A choice of what data to display and how should be made for each information-seeking context.

(3.2) Show reasons for actions. (See also (5) Structure).

Show where the user came from when he is at a place, how he got there, why he is there.

Show how the searcher got to an event/document and what he did to it.

(3.3) Show printing, saving, copying, sharing, and other previous use on documents.

(3.4) Show unexamined documents from result lists.

(4) Means of expression

(4.1) Use specific terms to label history items.

(4.2) Graphically differentiate general classes of items in display (e.g. queries in red or different icon).

(5) Structure

(5.1) Present structure graphically or through structured text (e.g. indent).

(5.2) Give the user a choice between graphical and text display.

(5.3) Show relationships between actions and documents, specifically show query-document pairs.

(5.4) Show duplicate or similar actions, repetitions of documents, query variations.

(5.5) Display digressions separately, show path with branches, show tangents (e.g. result list tangents, show progress on results list).

(5.6) Overlay of browse/search structure w/ information source structure, for example, showing the user's path within the graph of the hypertext structure.

Table 14. Search history display design guidelines.

Search history displays can range from a simple list of search steps to an interactive manipulable display of results and steps. Presenting the search history to the user as it is being built makes the user aware of the recording and creates an opportunity for her to stay in control of it. In addition it provides added functions to search systems, such as easy redo and undo functions. This section discusses guidelines for displaying search history information, the next section describes tools for managing search histories.

The guidelines in Table 14 were created based on the data collection in Phases One and Two. They are discussed below one by one in more detail.

(1) Align search history with user tasks.

(1.1) Different views of history needed for different purposes.

Providing different views of the search history for different user tasks is important. While the temporal display can facilitate navigating back to previous actions, organizing the history and results by topics supports results gathering and information use. Allowing the user to select among multiple views of the history puts the user in charge and allows him to adjust the display to support the task at hand. Flexible manipulation tools also facilitate viewing options, these are described in section 7.3.1.2.

(1.2) Parallel search histories for parallel tasks.

Attorneys often work on several tasks at the same time and would like to have access to multiple search trails. Keeping **separate tracks of** these **tasks** is important for supporting task management. Having several search histories open with easy switching between the active search trail is another important feature.

(2) Screen real estate, previews, overviews, and zooming.

(2.1) Do not make the history display rule the search system screen.

When presenting search histories, **screen real estate is of outmost importance**. Search histories can get very large in a brief time period, and it is impossible to present all steps and information on a portion of the user's screen. It is important to keep the goal of supporting information seeking in mind, the history display should be small in relation to the search system display, it is there to support searching and not take away space and user attention from searching.

(2.2) Provide easily accessible previews of history items.

Representations of history items should **present and/or link to both generic and specific information** about the history items they refer to. The class of the item, such as query, document, text section can be represented through graphical features such as icons or color as these can be described through a limited number of categories. Specific information such as query terms and document titles require text. Information that is not shown in the representation itself should be available through **easy previewing options**.

A preview shows some of the item's attributes, such as query terms, the database the query was run in, or the date and time of actions. Providing previews of history items in the condensed history display helps searchers decide where to go back and can encourage the reuse of older information. The importance of the various attributes and when and where they should be shown depends on the user and the task. Some of the attributes can be represented in overviews and history item views, while others only emerge with previews or through accessing the previous system state. Previews need to be accessible easily through the history display. Thus only representing history items through high-level surrogates with easily accessible previews can help save screen space.

(2.3) Provide overviews of action sets and sessions, based on chunking of steps along with the ability to zoom in on parts of the history.

Providing high-level overviews of the search history and document sets is important, as it allows the searcher to understand characteristics of a larger unit of information at one glance without having to scroll or zoom. An overview shows high-level abstractions of many items in one display. Aggregating steps, chunking the log into phases or tasks can help with this problem. Providing an overview of older history that the user does not need in detail at the moment can be helpful as well. Chunks in search logs can be created based on different queries, searches in different databases, or various topics explored. High-level overviews of the search history should be complemented with **zooming tools** that enable searchers to select a section of the search history and “zoom in” on it to see details; this is particularly important for the older parts of the search history. Providing overviews can also help the user in quickly identifying the sections of the log that is of interest. This way the user can navigate more quickly to the section of the log of most interest to her. Providing the overview can also be helpful in serving as a search map and orienting the searcher as to where she has been and where she needs to go. It can keep her from going off on tangents and finding her way back to the main sweep of the search.

(2.4) History item icons should link back to the event they represent. Make functions with specific content available.

History item icons should link back to the event they represent. Clickable lines representing queries and other actions, as implemented in the Westlaw search history display, proved to be successful, even though in Westlaw they disappeared when the searcher clicked on them to return to the action. History items in a display should take the searcher back to the system state that they represent with options to have new actions from that point or replay the previous process. This also achieves the goal of providing certain system function with specific content already filled in, such as accessing the search screen with specific search terms. Searchers should also be able to repeat actions with different attribute, such as rerun queries in new databases or edit queries.

(3) Type of data shown.

(3.1) Types of data to save were described in Chapter 6. A choice of what data to display and how should be made for each information-seeking context.

From the myriad of **history information** that can be potentially recorded from searching for information in computerized systems, a **selection** should be displayed to the user. The selection should be made on the information-seeking context: user characteristics and tasks, and system characteristics and tasks.

(3.2) Show reasons for actions. (See also structure).

Show where the user came from when he is at a place, how he got there, why he is there.

Show how the searcher got to an event/document and what was done to it.

Presenting why an action was executed is very important in enabling the searcher to understand the process. This information can be expressed by the context, or the searcher can explicitly enter this information. Presenting how a certain system state was reached can help with explaining why the action was taken, thus displaying where the user came from can help with this goal. When presenting documents, the routes by which they have been reached can also be presented along with the document.

In addition to displaying the context of the history item at the time of its occurrence, its relationships to preceding actions, with special emphasis on the ones closest to the history item is important. As a general rule, it was suggested that in case of every unit of information both the **action that led to it and the action that was taken on it** should be presented. A clear distinction between these two types of information is important. If a case was Shepardized, it should display that information, along with a link to the results of this action. If part of a case was copied to a document, it should reflect the copy action, and the same is true for emailing, printing, or saving documents. In this sense, the history item should reflect wear and tear as history-enriched digital objects. Displaying how the searcher arrived at a point in history can also help the searcher understand why a certain step was taken. Relationships between history items can express this, just as the actions that happened next.

(3.3) Show printing, saving, copying, sharing, and other previous use on documents.

Printing and saving are often not **recorded** in current search history systems, although these are very important events for the searcher. Both printing and saving mean the a document was judged to be relevant enough to keep for further use. Saving these actions in search history records is important, just as saving them on the history items they referred to. When the searcher views a document in the context of the search history, printing and saving can be displayed as part of the document representation, or a separate action following the document representation and linked to it. Showing other previous use on documents and other objects is also useful. Sharing, linking to, citing from are all important actions that characterize the document and will help searchers find a continuity of their actions.

(3.4) Show unexamined documents from result lists.

Filtering out documents from the result list based on their previous history with the searcher support the searcher in interpreting and managing the result list. Whether new or previously seen documents are marked depends on the system context: users, tasks, and the domain.

(4) Means of expression

(4.1) Use specific terms to label history items.

Using specific terms in history displays is very important in order to best support users. Generic terms such as "Query" or "Document" are not distinctive enough, using query terms or document titles serve better in these displays.

(4.2) Graphically differentiate general classes of items in display (e.g. queries in red or different icon).

While using generic terms in not helpful, informing the user about the type of item is needed. Presenting the category of items through graphical means is one solution for this problem. Each icon can represent a class of history items.

(5) Structure

(5.1) Present structure graphically or through structured text (e.g. indent).

Even if the search history is not presented in the context of the system structure, it is important to provide some kind of **structure in the history presentation**. The structure helps searchers interpret the search history by assigning units to the flow of events and defining some of the relationships between history items. The structure is determined by the order of the actions and objects, their types, and the relationships of the types to each other. This structure can be represented graphically, or through text displays with indentations or other formatting tools.

(5.2) Give the user a choice between graphical and text display.

Preference for graphical or textual structured displays was mixed in the study. Further examination of the question is needed to find out what the factors are influencing the choice for representation. The information-seeking context can influence this decision.

(5.3) Show relationships between actions and documents, specifically show query-document pairs.

In addition to presenting sufficient information about a history item, **its relationships to other history items and to the system** are part of the context and also very important for searchers to understand what happened, how they arrived at a history item and why. Presenting more of the context helps interpretation and is also beneficial when the searcher return to that system state. Presenting the relationships of the history items can be achieved through graphical representations or structured text.

An important attribute pair to include in search history displays is the **query-result set combination.** In addition, the number of results in the set are important for searchers as these numbers express a lot of information about the structure of the information source, topics and so on. For a full list of attributes of objects and action, see the chapters on search history use and search history data. Many attributes of objects and actions were described in the Search history data chapter, various combinations of these should be displayed along with history items depending on user tasks.

(5.4) Show duplicate or similar actions, repetitions of documents, query variations.

Showing similarities between objects and actions is important in **filtering out recurring documents or repeated action**. At each new action or object the system can compare it to the logs and signal to the user if a similar item is found in the log. Showing duplicate actions or documents in a search history is very important and can be achieved through comparing all new actions and objects to previous history. If identical or similar actions or objects are found, the searcher should be warned about the similarity in the search history display.

(5.5) Display digressions separately, show path with branches, show tangents (e.g. result list tangents, show progress on results list).

Presenting tangents as separate unit branches of the search history can help in the interpretation of events. Checking the Keycite history of a case in a results set is such a branch that the user takes from his major trail of actions: posing queries to the system and browsing the result sets. Marking the keycite history action as a digression helps with presenting the structure of the search history. The history display can present a tangent with a differentiating color or other feature. User marking or an analysis of user actions can delegate branches as tangents.

(5.6) Overlay of browse/search structure w/ information source structure

Overlaying the search history information on the structure of the information source(s) would help searchers understand their actions better. For example, presenting a user's browsing path over a website structure can help users by providing system context in to addition to the browsing path. The Ariadne interface discussed earlier achieved this through defining three levels of the systems and presenting search actions related to those three levels. Displaying history item classes helps this with simple system structures where clear classes of history items can be defined and linked to system components, such as key numbers or queries. With complex, less standardized hypertext systems this may be a challenge. Presenting the search history over the system structure also displays user position in a system environment which helps with the reuse of the action and also with the interpretation as it includes more from the context of the step.

7.3.1.2 Search history manipulation

Searchers want to save session histories for many different reasons, thus, basic storing and sharing capabilities need to be provided. Storing involves creating a new search history, starting recording, saving the search history and opening an old search history. Stopping or pausing recording and deleting search records is just as important for searchers' privacy and control. After the search is finished, the searcher should be able to open and manipulate the search history. Both while and after searching, reordering and searching search histories should be an option.

Sharing often involves some kind of user annotation described below to prepare the record for sharing and then the physical acts of sharing such as posting to a common database, emailing or printing the record including annotations.

Personalizing the search history record through annotation tools greatly enhance its usability. Three tools in particular were discussed in the design sessions.

(1) Annotating through localized text notes can help later interpretation.

(2) Marking and highlighting draws attention to certain parts of the record.

(3) Grouping is another option for managing search histories, it leads to the integration of the search history display with the organized collection tool and the scratchpad tool.

Planning as a function also leads to integration with other history tools, although it can also be provided through the search history display by devoting a portion of the time line to future actions where searchers can create reminders. The functions are summarized in Table 15.

The future of search history tools points in the direction of integration along several dimensions. Integrating different databases and information sources in one search history is one important dimension. Another is between user tasks using other applications and searching. Integration the search history display with the two other history-based searching tools is also very important in making the history display useful for searchers.

Objects:	**Actions:**
Search history	Create
Section of search history	Start recording
	Pause recording
	Stop recording
	Save
	Delete
	Open
	Close
	Manipulate the search history
	Search
	Reorder
	Annotate
	Mark and highlight
	Bookmark
	Group or cluster
	Plan
	Email
	Share
	Print

Table 15. Objects and actions in search history manipulation interfaces.

Design guidelines:

Recreate context with all or most aspects of information and situation

Make actions repeatable with different attributes

Integrate multiple databases, systems, sources in search history

Closely integrate the display with the collection and the scratchpad tools

7.3.2 History-based scratchpad tool

In searching there are many types of information that must be kept in memory for the next step; without computer support participants in the study recorded these in handwritten notes or by copying and pasting information into a word processing application. This observation led to the conclusion that a scratchpad tool is necessary where users can record information that they do not want to memorize, cannot keep in working memory long enough, but will need in the next step or screen.

The second interface tool suggested based on the data collected resembles the information retrieval whiteboard described by Spink and Goodrum (1996). It is termed "scratchpad" in this dissertation, as it allows the user a flexible space where various search support tasks can be carried out: it allows users to take notes, create knowledge models, write new documents, manage information-seeking tasks, plan and create reminders and save and share the results of all these actions. This tool takes search histories a step further: it attempts to help the searcher record thinking and information use in addition to just searching for information. It uses this history in helping searchers integrate new knowledge into old knowledge structures and create new products and documents from this process. Spink et al., in their study on reference librarians' searching notes found that searchers use many graphical elements in their notes. The roles of graphical markings are many, from creating models to emphasis and showing relationships. Graphical marking tool should also be available in the scratchpad tool.

Spink et al., in their study on reference librarians' searching notes found that searchers use many graphical elements in their notes. Some of these graphical elements serve to represent relationships and groupings. The scratchpad tool can support both of these functions, as illustrated by the list of objects and actions in 16.

7.3.2.1 Interpretation

Objects:	**Actions:**
Notes	Create
Knowledge models	Link
Search history	Mark
Document	Highlight
Concept	
Queries	

Table 16. Objects and actions in scratchpad tools.

The scratchpad tool helps searchers interpret new information found and relate it to old knowledge and thus learn through note-taking, knowledge modeling, and links set up between notes and knowledge models and the search results and search histories.

Tools to create graphical knowledge models to represent relationships of issues and the structure of a conceptual space are needed to support interpretation. A general diagram tool with "concept holders" and relationship markers can satisfy this requirement. The searcher should be able to copy and paste search history and result elements into these knowledge model, preserving the links between source and destination as described earlier. Graphical drawing, marking and highlighting tools should also be available to help the process of modeling. In addition to independent models, tools to create structures can be applied to the search history display. Providing a structured display was one of the recommendations made for search history displays. Allowing the user to reorganize this structure can be part of the search history manipulation functions. Another application of structuring tools is in case of the results storage. Storage areas need a structure to store documents and notes, this structure usually mirrors the user's task and knowledge structures.

It is important to note that all the lower level functions such as copy, save, and so on still apply here, but as they were covered earlier, they are not discussed again. The objects and actions necessary to provide the desired functionality are listed in Table 16.

7.3.2.2 Information use and document writing

Objects:	**Actions:**
Document outline	Create
Document text	Edit/Modify
References	Copy
Knowledge models	Convert to text outlines

Table 17. Objects and actions in information use and document writing.

The next information use step in attorneys' work is often the creation of a new document, brief or memo. The scratchpad tool should allow for word processing in close relationship to information seeking and interpretation. Earlier user notes should easily convert into the outline of a new document, preserving the links between source documents and designations in the new documents. Knowledge models should also transfer to text and create new document structures.

7.3.2.3 Task management

The scratchpad tool can also support the management of the information-seeking task. In collaborative environments it can help with task delegation and information sharing through saving and sharing information recorded in the scratchpad along with the search history. Incorporating a 'task' object into the scratchpad interface in the form of checklists that can be assigned top various team members, facilitates collaboration. Annotated search histories with tasks attached can be saved and sent to team members as a way to assign tasks to others.

Another important application area is the management of the searcher's own tasks. Many searchers start a topical exploration task by taking notes before going online to interact with information systems. These note can be taken in the scratchpad tool and should be easily converted into checklists of actions to complete. Information transfer between user notes and the search system should be smooth, the searcher should be able to drag a drop a part of his notes into the search system's query text box and create queries. This function can be complemented by automatic checking off on the checklist. Checklists and future actions can also be inspired by the new information the searcher finds through searching. Search result lists, queries and other search history elements should be easy to transfer into checklists of planned actions.

The searcher should be able to track his progress by checking off items on the checklist, or the system can automatically do this by matching the checklist item to the action in the search system. The checklist should be incorporated in the time line, the search history can extend into the future, and searchers should be able to click on future actions on the time line and cause the system to execute the action. This monitoring of the progress, along with results collection, can support the evaluation of the search and help searchers plan future action.

As task checklists are created, these can be automatically added to the "Future" section of the search history display, thus allowing for an easily accessible, small display of reminders incorporated in the search history display. As they are executed, they can transfer to the search history "Past" section. Larger scale planning of projects can also be accomplished with the help of these tools.

Objects:	**Actions:**
Checklist items = Task	Coordinate among attorneys
Checklist	Create (checklists):
Reminder	
	From scratch
	From user notes
	From search history or results
	Plan and organize
	Create

Table 18. Objects and actions in task management.

7.3.3 History-based result gathering tool

The organized gathering tool is another function that supports transfer from information seeking to information use. While the scratchpad tool enhances sense making through user notes and knowledge modeling and helps information seeking through task management functions, the organized collection tool helps searchers to collect results in an easily reusable form. The first step of this task is the management of result sets, which can be implemented through improvements to the search system or as an add-on within the organized collection tool. The next step is the collecting and saving of search results along with relevance decisions and user annotation to facilitate the later use of the information collected.

The scratchpad tool and the organized collection tool should be tightly integrated to ensure good integration between user notes, document outlines and the information collected.

7.3.3.1 Search result manipulation

Displaying search results can build on records of previous results by filtering out and marking results that have appeared before. The local systems can flag cases seen before and point these out to the user to support processing results.

As described in Chapter 5, end users often described a process of manipulating the results after identifying a result set of interest through searching. The actions involved comparing and combining result sets, searching within result sets, and annotating and marking results. Comparisons, combinations and searches within result sets can all be accomplished through complex Boolean queries, but end users may not have the expertise and felt more comfortable with directly manipulating the result sets. Objects and actions for search results manipulation are summarized in Table 19.

Objects:	**Actions:**
Query	Compare
Result	Combine
Document	Search within
	Show new items in result set (feature)

Table 19. Objects and actions in search results manipulation tools.

7.3.3.2 Search result selection

In most online search systems selection is a binary one-time choice. In online shopping there are examples of more flexible shopping cart collection tools where shoppers can stretch out purchasing decisions over time by marking items of potential interest and can even create temporary storage areas for items of interest, such as wish lists. In search systems, users should be able to save sets of documents for later review as they may seem promising, but not strong enough for reviewing them while searching. Weak relevance is defined as relevance judgments made about documents but not fully committing to the relevance of the document. If the user sees something that might be relevant, and marks it to come back to it later for a more in-depth investigation. Searchers need guidance to come back to these documents without forgetting about them. Marking these documents serves as a buffer to store them until later.

Relevance decisions in information seeking are influenced by many different factors and they can change dynamically as the knowledge of the searcher is updated through finding new information. As described in earlier sections, documents that may seem relevant at the beginning of the search can become irrelevant by the end of the search, and via versa, documents ignored in the early stages of the search may become important later. In the latter cases, searchers should be able to retrieve these documents from the search history and transfer them to the collection tool.

Although the decisions to save that are made in the selection phase may not be final, even temporary relevance judgments should be recorded along with the documents. The relevance judgment can be recorded through a form, annotation and notes, and highlights and graphical marks. A short reminder of the type of relevance can also help when searchers forget why they saved or printed a particular case. When entering relevance judgments, a pop-up window with a relevance template can help users in relating the information found to a problem or task. The relevance template should have the following items:

(1) Space for user notes

(2) Checkboxes with the various attributes of the document to mark which attribute made it relevant

(3) Tools to rate the degree of relevance on a scale

Searchers make relevance judgments based on many different attributes, depending on the task. Saving the attribute that the decision was based on can answer the question of why it was relevant. This selection is hard to detect automatically, the user has to select it if he wishes. In the pop-up window described above, a list of attributes can be displayed, from which the user can select the decisive ones. The attributes presented can be taken from the document.

The relevance decision information will be used when applying the information found to a task, e.g. writing a document. This information may include pointers to how the information is supposed to get used. By showing why a case or document was relevant at the time of the decision the user's task of interpreting and integrating documents at a later time is made easier by reminding them of the previous interpretation. When making a relevance decision and saving results, these can be linked to the task they will be used for right away. One case can be relevant for many different topics or tasks. Users should be able to note that on the case. Lin's work proposes a system to do customized indexing of documents found. His tool provides functionality to assign documents found to many interests or tasks of the user.

The recording of relevance information should be optional for the user, as it helps with some tasks and is unnecessary with others. Although the relevance decision may change, it is important to record reasons for the decision at the time it is made. In addition to relevance annotations, users should be able to create generic notes and annotations linked to the documents selected. Table 20 summarizes actions and objects for search results selection.

Objects:	**Actions:**
Relevance judgment	Two-step saving of documents with temporary storage area
Attribute	Mark
Degree of relevance	Annotate
	Transfer (documents) from search history

Table 20. Objects and actions in search results selection tools.

7.3.3.3 Saving and organizing search results

After a selection decision is finalized, the selected document is saved for later use. Table 21 describes objects and actions involved in saving and organizing search results. When saving and storing search results, attorneys organize them by several different attributes. Topic, client ID and task are the most important attributes to organize by. Organizing saved items by these attributes can be implemented through different solutions: assigning items to categories, or assigning index terms to items. In either case, the searcher should be able to display the items by the various attributes. The searchers should be able to index the items by multiple values of the same attribute if necessary.

Creating the structure of topics is a demanding task, but it is one that attorneys are willing to take on, even in less flexible paper filing systems. As described in Section 7.3.2.1, the scratchpad tool envisioned here provides support for creating the topic, client ID and task categories.

Attaching history items to one or more tasks is also important. Saving relevance judgment information was discussed in the section above but should also apply to search results saved for the long term. Annotation and marking tools are also necessary here.

Whether notes, results and search histories are saved or not is influenced by how easy it is to save them. As expert searchers often take notes of steps and results, and suggest it to inexperienced information seekers, it is a behavior to be encouraged while searching. If it easy to do, information seekers may be more inclined to save information. Information in various formats requires different techniques for saving.

Simplifying the process for all formats is important.

The presence of the collection tool on the screen along with the search systems helps users focus on selected topic areas represented by topical folders or indexing terms in the collection tool. It also makes it easier for searchers to constantly evaluate their collection and reevaluate their needs. Providing the same collection tool across various databases helps the task of integrating results from multiple sources by crating a permanent and accumulative collection of results with shared tools.

When saving results of information seeking, it is very important to preserve the link between the documents and the source: the action in the search history that led to the results, for example the query and the database. The searcher should be able to traverse back to the event from the result saved.

When searching for information, users may find information that is not relevant to their current task, but is useful for a different task. The system should provide an option for searchers to save these serendipitously found results linked to the task it is relevant for and create a reminder next time the task is accessed.

Objects:	**Actions:**
Documents	Save
Organization structure	Organize
Annotations linked to results	Create organizational structure
Source	Integrate results from multiple sources
Serendipitously found information	

Table 21. Objects and actions in tools to support saving and organizing search results.

7.3.3.4 Document manipulation

After documents have been saved, searchers need to manage and manipulate the documents. Searching within a document is one of the important functions needed. In order to allow the user to accomplish this, the full text of the document needs to be saved. Annotating and marking within documents is another important need described by study participants. Linking user notes to specific sections inside the document is very important in using information, sticky notes attachable to page sections are one example of this tools. Similarly linked graphical marks are also necessary.

Integrating search result documents with user-created documents is a requirement many searchers in the study described. Being able to search across the saved search results and the user's own documents is a powerful tool. Table 22 provides a summary of objects and actions.

Objects:	**Actions:**
Documents	Search within
	Integrate with searcher's own documents
	Mark
	Highlight
	Link to specific part of a document

Table 22. Objects and actions in document manipulation tools.

7.3.3.5 Integration

One of the roles of the collection tool is to help searchers integrate information found into their information use tasks. Many of the functions described below support integration. The collection tool, by allowing the searcher to link results found to documents to be written or to other tasks, can prepare the results for reuse. Creating annotations attached to documents can also help the same purpose. The scratchpad tool also includes functions to help integration.

Objects:	**Actions:**
Tasks	Integrate
Search history	Use
Results	Reuse
Searcher's own documents	Write
Document building blocks	
Personal libraries	

Table 23. Objects and actions in integration.

7.4 General principles emerging from the data

The following general design guidelines emerged throughout the study and can benefit interface design. As these guideline are more generic, the objects and actions are presented as pairs. These guidelines are valid for all three interface tools described earlier (search history display, scratchpad, results collection). While these tools described specific function of user interfaces, the following two design recommendation are more general and to be applied in all three tools and other interface designs. These recommendations do not necessarily build on recording search histories, but create connections between user steps separated by time in computer applications.

7.4.1 Shorten distance between interaction events

Shortening the distance between interaction events enables searchers to view and return to previous events more easily. Allowing easy access to previous events (through such tools as search history displays, and action comparisons) gives users a wider view of their actions; providing this information makes it easier for users to remember what they have done previously. The various user interface (UI) tools described below provide support for this guideline through different techniques.

In addition to easy access to logs, closer relationship between steps can be created through comparisons between actions, queries and documents. When a query is entered, similar previous queries can be displayed in a drop-down list; results documents can display query terms from several queries back through color coding, re-emerging results documents can show repetition through icons or color coding. Similar actions or results in temporal search history representations can express this through a representation of the relationships. Comparing items in a search history log is a way to bring events closer together.

One possible feature is an autocomplete function: when a user enters a query word, the system can show previous queries that contained that word, thus making similar earlier functions more easily accessible.

Object action pairs:

- Easy access to log/previous steps
- Highlight search terms in result documents
- Compare steps

7.4.2 Easy reuse of previous information

Easy transfer of information from one action to the next preserving source information is another important feature that also supports shortening the distances between actions by creating linkages through reused data. Currently, copy and paste functions are often used to achieve reuse of previous information. Copy and paste functions serve as multifaceted tools in search history and results management. Copy and paste are very simple to use and understand in graphical user interface environments, they are very versatile and flexible. These characteristics made these functions widely used among searchers for various history-related tasks.

Frequently used copy and paste transfers, for example adding a keyword to a query, should be accomplished with a single or double click rather than dragging.

Keywords found in results are often reused in new queries, the transfer from results to queries should be easy. Citations, links, names of people, places or sources may also be reused in further queries. Keywords may not only be reused in queries, but also in interpretations, documents written, etc.

Reuse can come from fast re-creation, not just recording. In a flat-fee scenario, the searcher may simply retype a query and this may be easier and less time-consuming than pulling it up from a log. In this case, the system can match it to earlier entries and make those available.

Reuse is facilitated by easy moving of items between interface areas. One way to move items is to leave a copy of the item in its original location and paste a copy of it in a new location. A complete move would remove the items from its original location. It is important for the item to 'remember' the source it came from so that its move can be retraced if needed.

Object action pairs:

- Use results information in new queries
- Automate recurring searches
- Edit query
- Rerun query
- Use information found in query
- Reuse keywords in new queries
- Update search
- Search within result sets
- Show on document that it was copied from
- Show source when copied text

7.4.3 Importance of design features

While all the functions and features described are important based on the data collected, the realities of system and interface design call for a prioritization of the features. In light of the actual user groups and context for system design, decisions need to be made about the order in which features should be considered. The user group in this dissertation is made up of two distinctly different groups in two slightly different contexts: reference librarians in law libraries and practicing attorneys working from their personal workstations. Design recommendations for these two groups are made separately, as their needs differ.

However, the first priority and basic interface tool to be provided is the same for both groups. Recording search histories creates the foundation for all of the tools described in this chapter. Once automatic recording of the interaction history is provided, displaying this information to users can be more easily implemented. A search history display without some manipulation tools and access to the earlier states represented is not predicted to make a significant difference in the usability of the interface. When a search history display is provided, notes, annotations, planning functions should also be provided. Another important feature to include is making the search history icons lead back to the previous system state they represent.

As mentioned above, the requirements of the two user groups examined were different. For the search intermediary tasks of law librarians, the organized results collection tool plays a more important role. Search intermediaries interpret less of the information, and focus more on collecting a satisfactory set for their patrons. In this case, the collection tool can help law librarians create deliverable sets organized according to the needs of the patron.

For the end-user attorneys, the scratchpad tool can provide the necessary support in interpreting new information and preparing it for reuse in their tasks. While this user group is also interested in collecting and organizing documents, their main goal is to extract and use information from the results of the searches. The features of the scratchpad tool can support their needs in light of this overall goal through providing note-taking and annotation tools and other links between the information found and their final products.

Designers will have to make the ultimate decisions about the priority of the features for each systems based on the user group, tasks, and context variables at hand. While the generic guidelines described here can help with this task, a user needs assessment of the concerned user groups cannot be substituted by these recommendations. The characteristics and tasks of the users, their organizational context, and other factors will have a strong influence on what features should be implemented first.

8 Conclusions and implications

Recording search histories, presenting them to the searcher, and using them to build additional interface tools offer many opportunities for supporting user tasks in information seeking and use. This study investigated the use of search history information in legal information seeking. Qualitative methods uncovered how attorneys and law librarians used their memory and external memory aids while searching for information, and in transferring information to use. The findings form the basis for interface design recommendations.

The results of the study have shown that legal information seekers use search histories. If the system does not provide adequate histories, users create pieces of it to support their memory and help them achieve their goals. User observations and interviews revealed that search histories support many different user tasks in information seeking, from task management through learning to collaboration. Different types of search history data and interface tools are needed to support these various tasks and to provide integration across tasks. Interface tools based on an automatic recording of interaction events, supplemented by user-entered data, such as notes and annotations, are needed. These tools allow searchers to record and manage various types of information across the sub-processes of information seeking and use, and help them reuse the information recorded.

Search histories as external memory support tools proved to be useful for information seekers in many task areas. Chapter 5 described these areas along with some implications for interface design. Chapter 6 defined the types of search history information needed to support these tasks and explored topics related to the saving and management of this information. Chapter 7 built on Chapters 5 and 6 in bringing together interface design guidelines to support user tasks by applying appropriate search history information to the user interface.

The understanding of user tasks and needs is the first step in the design of usable computer interfaces. User observations and interviews, and an analysis of existing literature on information-seeking behavior informed the design of interface tools. This assessment of user needs involved an investigation of how histories were used in searching; what elements of search sessions were important for the user to be saved; and what interface techniques were most appropriate to support effective use of histories and successful information seeking.

8.1 Search history framework

The overarching goal of the dissertation research was to develop a search history framework defining and describing search histories, laying out the territory and dimensions of search histories, their uses, definitions of data to be included, and interface design suggestions. The framework identifies user tasks that search histories can support, factors that influence search history use, search history data to record, and user interface tools that can be built based on recorded search histories. This framework guides students of information-seeking behavior and system designers in identifying potential areas of application for search histories in their context, selecting search history data they need to record to support these areas, defining appropriate management tools for this information, and potential interface tools they can design and build.

The final version (Version 2) of the framework is presented in Appendix A. It has been developed through three iterations in the dissertation. Version 0 was created based on the researcher's personal experience, the results of the literature review, and the outcomes of the opportunistic data collection efforts (reference librarian interviews and usability testing videos). The results of the attorney and law librarian observations and interviews updated the framework to Version 1. This version provided the basis for developing user interface designs, and results from the participatory design session were used to update the framework to Version 2. As described earlier, the framework is organized into the following facets:

(1) Scope of history

(2) Context of history

(3) Search history data

(4) Search history and result management

(5) Search history use

(6) Design features

The scope of the history sets the dimensions along which search histories can be defined. The items in this category can lead the definition of the search history system at the beginning of a design process. A certain value along the dimensions will clearly describe the type of information recorded in the search history. The dimensions for defining the scope are: time span, task span, system/source span, user span, and data source. Time span describes the temporal coverage of data recorded, history can be kept and provided from the current sessions, from several session before, the user's life history with a certain system, and they can be extended into the future by including user planning. Similarly, task span refers to the types of tasks included in the history, information-seeking tasks only, or information use and other user tasks as well. Search histories can contain information from several or just one system, depending on how many systems the user searched, or how many systems the history spans. The number of users dimension distinguishes between individual and shared search histories are. Finally, history data can be recorded automatically by the system, or manually by the user. This facet is described in detail in Chapter 3.

The context of the history describes the task domain and characteristics of the domain that will affect the use of search histories. It considers the nature of the task domain, characteristics of organizational context and information access variables. User and task factors are also defined here. This facet is described in Chapter 3.

The search history data category defines the types of information to be saved and helps system designers select information to record in their own systems. The list does not aim to be complete and is reflective of the legal information domain and the needs of users in the study. The search history and results management category describes user needs related to the management of search history data and of data resulting from the search, which are closely intertwined. This facet is described in Chapter 6 along with search history data.

The search history use section is the largest in the framework. It describes the potential user task areas where search histories can be helpful. In studying information-seeking behavior or designing system interfaces, others can refer to this section and identify application areas relevant to their context. The range of areas is based on legal information seeking. On the one hand, this limits the variety of task areas to those relevant to legal information users. On the other hand, it provides a detailed picture of what tasks legal information users have in terms of information seeking. This facet of the framework is described in Chapter 5.

The design features facet reviews possibilities for user interface design solutions building on the recording of search histories. As the other facets, this category describes a set of potential features: the list should not be considered as complete or as a prescription for all systems. It provides ideas for designers that they can pick and choose from. The application of certain features is strongly influenced by the users, tasks, and context of the system. This facet is discussed in Chapter 7.

The goal of this framework was to define areas of concern in relation to search histories for this research and for future researchers studying the topic. It defines the various terms and their relationships to each other in order to systematically review the issues in connection with search histories. A more detailed discussion of the different facets of the framework is found in Chapters 3, 5, 6, and 7.

8.2 Summary of results

The results of the study suggest that recording search history information is very important for successful information-seeking interfaces. Searchers use history information extensively through system tools and create individual external memory aids if the system does not provide them. The appropriate history tools should be selected based on user tasks and context. Interactions histories in computer applications can support many different tasks in various domains. This dissertation examined searching for and using information as a particular task domain. Some of the results are specific to this task domain, while others can be generalized to other systems and user tasks.

Search history information can be recorded by the system so that searchers do not have to remember it. It is always available to the searcher through user interface tools; the searcher can easily return to previous actions or reuse previous information unchanged or with modifications. Recorded search histories support information seeking and use through providing tools for task planning and management, including monitoring progress, evaluating actions and creating reminders. In navigation, the search history record helps searchers orient themselves and find their way around the system. When returning to a search task after an interruption, search histories help searchers continue the information-seeking task. Learning, developing mental models through interaction with newly found information is carried out all through searching. Providing history information helps with the accumulation and recording of this information, along with preparing it for reuse. Search results collection and management also benefit from history-based tools that provide the continuity of collection across many steps of searching and various sources, and allow the searcher to organize, annotate, and search the results.

Providing interaction history information in a computer application helps widen the user's context and vision area through displaying previous actions and events an creating space for planning for the future. User and system actions separated by time are brought closer to one another through the continuity of time ordered displays of events. Providing this extended context, computer application more closely approximate physical environments where history and wear is more visible and accessible to users.

Accumulation of user actions is natural in physical environments, while in digital environments wear and tear may not be obvious on objects, and the lack of physical landmarks can decrease the user’s ability to remember previous actions. This type of history information is related to the object, including the history of the object within its representation. In addition to object history, temporal process histories are needed to support information seeking. By providing a tools that display temporal accumulation of events and information, the searcher’s mental model building and learning can be advanced. Creating a process view of actions rather than just the current step combined with a couple of previous steps, the user can better understand the flow of actions and plan and execute his search.

8.2.1 Search history support for user tasks

Memory support

The major findings of the study show that legal information seekers rely on search histories as memory aids to support search system use, task management, information seeking and use, and in integration with other tasks and other people. The most elemental functions of automatically recorded search histories is memory support. By recording many of the actions and objects from the process, the system empowers the searcher to focus on higher level tasks than remembering details. In addition to automatic recording, the user can add more personalized information through notes and annotations, thus recording thoughts and ideas in addition to facts.

Search system use

The research presented in this dissertation concerned electronic information seeking. In this task context, the ability to use the search system is an important prerequisite to finding information successfully. Search histories can support users with this task. By marking places in the system, or seeking out specific actions in a search history, searchers can easily find functions in a search system. Navigating back to previous actions does not simply take users to various functions, but to these functions with specific content in the function, such as a query boxed filled in with a specific query. The data collected offered many opportunities for supporting the use of the search system by recording search history information.

Information seeking and results management

The process of information seeking involves many steps from entering queries to viewing and selecting results. Many of these steps can benefit from recorded search histories. Query formulation and reformulation are good candidates from the first half of the process, keeping track of what queries have been tried and what result sets they returned can help guide the user in future steps. Collecting and managing results can also be supported both by automatically recorded histories and allowing the user to manage these histories to select, organize, annotate, and save results.

Information use

Using the information found is the ultimate goal of information seeking. Recorded search histories can provide strong support in creating linkages between finding and using information through enabling the user to capture contextual information at the time of saving results. Noting why a certain document was saved and how it will be used can greatly enhance its use later on through reminding the user of his thinking at the time of selecting and saving the document. Creating annotations on documents, grouping results under the task they will be used for, or arranging them in future documents.

Task management

Information seeking is a complex problem-solving task. Searchers need to manage their tasks: plan for it, execute the steps, monitor progress, and evaluate the process and the end results. Tools that can support this task include note-taking features, to-do lists, checklists, and other planning tools, and displays of the process completed. Integration between the task management tools and the information retrieval system is very important in making these tools useful. Transferring information between the planning tools and the system where they are executed and automatically executing the actions is the input side of this requirement. Checking off completed actions automatically based on the record of the history represents the synchronizing the system output cycle to the user. Evaluating actions can be supported by a display of previous actions, comparisons across various actions can also support system learning.

Mental model building

Mental model building and learning is carried out all through information seeking. While search intermediaries may focus on learning about sources, search techniques, and types of information, end-user searchers, such as attorneys learn about the legal area. Mental model building is carried out in every search step, support tools should be present in each step.

Task integration

In addition to the searching tasks of individual users, the analysis of the collected data highlighted two other important use areas: integration with other tasks and integration with other people. In physical library environments, searching for information was geographically separated from the use of information: writing documents, informing colleagues and clients, and other tasks were completed in the attorney's office, while searching for information took place in the library. When attorneys search for information from their own personal computer, write documents, and communicate through email on the same computer, this separation has almost but disappeared; it is mostly represented by the separation of different computer applications. Creating histories across applications helps bridge these separations and provides the user with a task-centered process.

To be optimally effective in supporting a lawyer's work, information systems must mirror the holistic nature of this work. Rather than having separate applications for searching, managing personal research files, task scheduling, and personal information management, and writing, the lawyer should be supported by one integrated environment. Search histories can play a role in unifying this environment through bringing actions and data closer in time and providing a continuity across tasks. Within such an environment, the system should keep track of user actions and their results and use this expanded history to support the user's work.

Collaboration

Many users reported sharing search history records for the purposes of collaboration and coordination between team members, supervisors, assistants, and other coworkers. The area of using search and other activity histories for the purpose of collaboration has not been extensively explored in the dissertation, but promises potential avenues of further research, as discussed below.

8.2.2 Interface design recommendations

The interfaces designed in the framework of the dissertation were based on an extensive user needs assessment and later tested and further developed using a second set of methods, participatory design sessions. The Phase 1 user needs assessment highlighted search history use areas and enabled the researcher to design a preliminary set of interfaces, while the Phase 2 participatory design sessions helped test and further develop the designs, and ensure their appropriateness in supporting searchers' information seeking.

Three interface tools were suggested and described:

T1 direct search history displays and management tools,

T2 a scratchpad facility, and

T3 an organized results collection tool.

In addition, two general interface design guidelines have been proposed:

G1 shorten distance between interaction events,

G2 easy reuse of previous information.

Direct search history displays visualize search history information for the user and allow him to navigate back and forth along the time line. In addition to the display, several management tools were described, such as saving and deleting histories, repeating and modifying actions. The scratchpad tool proposed creates a space for the searcher to work with the information, create plans and reminders, and take notes. This tool is closely integrated with the organized collection tool, which helps searchers in collecting information from searches and preparing it for reuse in future tasks.

In the design of user interfaces for any computer applications, bringing interaction events closer together can help users better understand the process, take advantage of previous steps, and plan future actions. One way to achieve this is to enable users to easily recall and reuse previously seen information. By providing easy access to previous information through the search hsitory display, retrieval and reuse of this information is facilitated.

8.3 Future research

Development and testing of interface tools and their impact on information seeking

The interface tools proposed in the dissertation can be implemented and tested through more rigorous methods to examine wether the integration of search history interface tools will result in improved user performance and satisfaction. Examining the impact of providing search history tools is an even more interesting area. Whether the flexible availability of historical information changes the nature of information seeking can be examined if tools are available.

Specifically, the following questions need to be studied after the implementation of history tools:

(1) Do history-based interface tools help people find better information?

(2) Do they help them find this information faster?

(3) Do they help searchers in planning and managing their tasks?

(4) Does displaying the continuity of the process help people understand the process better?

(5) Does it help them learn about the topical area faster?

(6) Does it help them solve their problem better?

(7) Do these tools help searchers with their larger task?

Question 5 is especially interesting. Whether history tools help searchers learn about a topic or understand new information better can be tested through giving the same problem (e.g. understand a political situation in a country or a psychological problem) to two groups of users, providing one group with a normal display, while another a display with search history information integrated and measuring their performance on the original problem.

Another interesting and related question for future research is studying individual differences in using search histories and history-based tools. The data collected in the dissertation pointed in the direction that some cognitive characteristics of individuals influence the use of search histories, a more systematic examination of this questions remains for the future. An exploratory study of attorneys working in the same context is needed to identify what characteristics might influence usage patterns and then a more controlled experiment can be conducted comparing individuals with differing values on those characteristics.

Extending the scope of the study

Time Span					
Task Span		Current session	Several sessions	Lifetime	Future, planning
	Search	X	X		X
	Search & Info Use	X	X		X
	All Tasks				
User Span: Individual user/multiple users and collaboration					
System/Source Span: Individual/multiple sources					
Data source: automatically/manually recorded information					

Table 24. Scope of dissertation research.

As shown Table 24 and discussed earlier, the current study focused on search history on the individual users' searching tasks from the current session and several sessions back. Using search histories to support multiple tasks and multiple users emerged as important use areas for search histories, but have not been examined in detail. Integration across user tasks can be supported by recording history information in a common system and providing access to previous information regardless of the application it was recorded in. This support would shift the focus from tools to tasks, by allowing the user to select the units of activities across applications.

! What history information can and should be recorded across applications?

! How to display history information across applications?

! How would users use across-application histories?

! What history-based tools do users need across applications?

Individual task histories are often recorded for the purpose of sharing them with others. A common problem for organizational memory systems is the task of preserving and integrating individual experience and knowledge to build representations of the organizational memory. Automatically recording actions and allowing searchers to manage, organize, and annotate these records can provide potential outputs that can be shared with and used by others.

! How do searchers share search histories?

- Why do searchers share search histories?
- What tools do searchers need in order to create search histories that can be shared with and used by others?
- How can individual search histories be integrated into organization memory systems?

Finally, recording and preserving the user's life history poses a set of very interesting research questions in terms of data selection, archiving, and tools to manage life history data. The potential application of mobile devices can incorporate an even wider range of user tasks.

Extending the context of the study

The context of the dissertation research was the legal information field, the participants were attorneys and law librarians and the databases used were contained legal information. While the assumption was made in the beginning of the research that legal users can be especially good candidates to benefit from search histories because of the character of their work, as the search progressed it became more and more apparent that other topical domain and search tasks could also take advantage of history-based tools. Extending the examination of history tools to general Web searching is an interesting avenue of research, while focusing on other professionals such as medical or science users can enhance the findings of the dissertation research.

- Do Web searchers need history information to support their information seeking?
- What kind of history information and tools do Web searchers need?

Appendix A. Search history framework, Version 2. (Also served as Coding Protocol)

The framework presented here is a summary of the search history framework described in Chapters 5, 6, and 7. On the one hand, this summary is more comprehensive, as it includes more possible options, some of which may not be described in the chapters. On the other hand, the chapters are more detailed in describing each facet of the framework.

(1)	Scope of history
(1 1)	. Time span
(1 1 1)	. . Current session
(1 1 2)	. . Future - Planning
(1 1 3)	. . Immediate history
(1 1 4)	. . Lifetime history
(1 1 5)	. . Several sessions
(1 2)	. Task span
(1 2 1)	. . All activities
(1 2 2)	. . Integration of SH into personal DB
(1 2 3)	. . Search
(1 2 4)	. . Search and information use
(1 3)	. System/source span
(1 3 1)	. . Multiple systems or sources
(1 3 2)	. . Single system or source
(1 4)	. User span
(1 4 1)	. . Multiple users
(1 4 2)	. . Single user
(1 5)	. Data source
(1 5 1)	. . Automatically recorded data
(1 5 2)	. . User selection of automatically recorded data
(1 5 3)	. . Manually recorded data
(2)	Context
(2 1)	. Domain: Legal information seeking
(2 1 1)	. . Characteristics of legal information
(2 1 1 1)	. . . Context sensitive
(2 1 1 2)	. . . Interconnetcted
(2 1 1 3)	. . . Textual
(2 2 1)	. . Legal IS tactics
(2 2 1 1)	. . . Serendipity
(2 2 1 2)	. . . Trial&error searching

(2 2)	. Organizational context
(2 3)	. User characteristics
(2 3 1)	. . Attitude toward search history
(2 3 2)	. . Domain expertise
(2 3 2 1)	. . . Domain experienced
(2 3 2 2)	. . . Domain inexperienced
(2 3 2 3)	. . . Domain intermediate
(2 3 3)	. . Search intermediary
(2 3 4)	. . Searching expertise
(2 3 4 1)	. . . Searching experienced
(2 3 4 2)	. . . Searching inexperienced
(2 3 4 3)	. . . Searching intermediate
(2 3 5)	. . System expertise
(2 3 5 1)	. . . System experienced
(2 3 5 2)	. . . System inexperienced
(2 3 5 3)	. . . System intermediate
(2 3 6)	. . Type of work
(2 4)	. Task complexity
(2 5)	. Access variables
(2 5 1)	. . Price of online access
(3)	Search history data
(3 1)	. Search history in general
(3 1 1)	. . Overview of search actions
(3 1 2)	. . Overview of search results
(3 2)	. Session/housekeeping data
(3 2 1)	. . Session metadata
(3 2 1 1)	. . . Client ID
(3 2 1 2)	. . . Cost of session
(3 2 1 3)	. . . Length of session
(3 2 1 3 1)	. . . . User-created links
(3 2 1 3 2)	. . . Project
(3 2 2)	. . System settings
(3 3)	. Request data and query formulation data
(3 3 1)	. . Query data
(3 3 1 1)	. . . Fields
(3 3 1 2)	. . . Keywords
(3 3 1 3)	. . . Name of database selected
(3 3 1 4)	. . . Need-by date
(3 3 1 5)	. . . Previous information
(3 3 1 6)	. . . Query formulation

(3 3 1 6 1) . . . . Connectors
(3 3 1 6 2) . . . . Content
(3 3 1 6 2 1) . . . . . Numbers
(3 3 1 6 2 2) . . . . . Terms
(3 3 1 6 3) . . . . Fields
(3 3 1 7) . . . Query reformulation
(3 3 1 8) . . . Query statement
(3 3 1 9) . . . Request
(3 3 1 10) . . . Type of query
(3 3 2) . . Search navigation data
(3 3 2 1) . . Citation chaining
(3 3 2 2) . . Search phases
(3 3 2 3) . . Checking helpfulness of results
(3 3 2 4) . . Development of search strategy
(3 3 2 4 1) . . . Conceptual query formulation
(3 3 2 4 2) . . . Database selection
(3 3 2 4 3) . . . Database specific query formulation
(3 3 2 5) . . Editing search results
(3 3 2 6) . . Execution of search strategy
(3 3 2 7) . . Recognition and statement of need
(3 3 2 8) . . Review of search results
(3 4) . Search result data
(3 4 1) . . Existing links among documents
(3 4 1 1) . . . Anchored links
(3 4 1 2) . . . Customized clusters of documents
(3 4 1 3) . . . Link types
(3 4 1 3 1) . . . . Case to case
(3 4 1 3 2) . . . . Case to law
(3 4 1 3 3) . . . . Citation to current document
(3 4 1 3 4) . . . . Citation to other documents
(3 4 1 3 5) . . . . Law to case
(3 4 2) . . Individual document data
(3 4 2 1) . . . Document metadata
(3 4 2 1 1) . . . . Author of document
(3 4 2 1 2) . . . . Copyright information
(3 4 2 1 3) . . . . Court of decision
(3 4 2 1 3 1) . . . . . Jurisdiction
(3 4 2 1 3 2) . . . . . Level of court
(3 4 2 1 4) . . . . Date of decision
(3 4 2 1 5) . . . . Date of events - if case
(3 4 2 1 6) . . . . Events of case

(3 4 2 1 7) . . . . Format
(3 4 2 1 8) . . . . Language
(3 4 2 1 9) . . . . Length
(3 4 2 1 10) . . . . Parties
(3 4 2 1 11) . . . . Publication date of document
(3 4 2 1 12) . . . . Publisher
(3 4 2 1 13) . . . . Relevance judgment
(3 4 2 1 13 1) . . . . . Look-back relevant
(3 4 2 1 13 2) . . . . . Relevant for other topic
(3 4 2 1 13 3) . . . . . Weak relevance
(3 4 2 1 14) . . . . Ruling~holding of decision
(3 4 2 1 15) . . . . Topic of document
(3 4 2 2) . . . Own document
(3 4 2 3) . . . Part of document
(3 4 2 3 1) . . . . Abstract-Synopsis of document
(3 4 2 3 1 1) . . . . . Customized abstract
(3 4 2 3 1 2) . . . . . Headnotes of document
(3 4 2 3 2) . . . . Full text of document
(3 4 2 3 3) . . . . Key numbers assigned to document
(3 4 2 3 4) . . . . Selected sections of full text
(3 4 2 3 5) . . . . Title of document
(3 4 3) . . Result set data
(3 4 3 1) . . . Number of results for a query
(3 4 3 2) . . . Overview of search results
(3 4 3 3) . . . Result list
(3 5) . Metadata general
(3 5 1) . . Author
(3 5 2) . . Client
(3 5 3) . . Cost
(3 5 4) . . Project
(3 5 5) . . Title
(3 5 6) . . URL
(3 6) . Search result management data
(3 6 1) . . Data on sharing search results
(3 7) . User notes
(3 7 1) . . Length of notes
(3 7 1 1) . . . Long annotation
(3 7 1 2) . . . Medium annotation
(3 7 1 3) . . . Short annotation
(3 7 2) . . Subject of notes
(3 7 2 1) . . . History related user notes

(3 7 2 1 1)	. . . . Database characterization
(3 7 2 1 2)	. . . . Reminder for future search actions
(3 7 2 1 3)	. . . . Representation of mental model
(3 7 2 1 4)	. . . . User notes on queries
(3 7 2 1 4 1)	. . . . . Annotations on query elements
(3 7 2 1 4 2)	. . . . . General notes on the query
(3 7 2 1 4 3)	. . . . . Notes on keywords for later use
(3 7 2 1 5)	. . . . User notes on search results
(3 7 2 1 5 1)	. . . . . User notes on individual documents
(3 7 2 1 5 1 1)	. . . . . . User notes on document parts
(3 7 2 1 5 1 1 1)	. . . . Mar doc parts as irrelevant
(3 7 2 1 5 1 1 2)	. . . . Mark doc parts for reuse
(3 7 2 1 6)	. . . . User notes on SH representation
(3 7 2 2)	. . . History unrelated user notes
(3 7 2 2 1)	. . . . Reminder for future nonsearch action
(3 7 2 2 2)	. . . . Thinking about task issues
(3 7 2 3)	. . . User notes on notes
(3 8)	. System notes
(3 8 1)	. . System-created links
(3 9)	. Data on user actions
(3 9 1)	. . By type of action
(3 9 2)	. . Links among actions
(3 9 3)	. . Reasons for user actions
(3 9 4)	. . Sequence of user actions
(3 9 4 1)	. . . Important point in sequence
(3 9 5)	. . User action log data, metadata
(3 9 5 1)	. . . Duration of action
(3 9 5 2)	. . . Duration of session
(3 9 5 3)	. . . Name of searcher
(3 9 5 4)	. . . Number of steps
(3 9 5 5)	. . . Time of action
(3 10)	. Data selection issues
(3 10 1)	. . Data selection based on user needs
(3 11)	. Data storage issues
(3 11 1)	. . Duration of storage
(3 11 2)	. . Granularity of data storage
(3 12)	. Links
(4)	Search history and result management
(4 1)	. Actions on search history or results
(4 1 1)	. . Compare

(4 1 2)	. .	Copy
(4 1 3)	. .	Delete
(4 1 4)	. .	Look at
(4 1 5)	. .	Mark
(4 1 6)	. .	Print out
(4 1 7)	. .	Record
(4 1 8)	. .	Save
(4 1 9)	. .	Share
(4 1 10)	. .	Store
(4 1 11)	. .	Update
(4 2)	.	Finding data in search history
(4 2 1)	. .	Navigate search histories
(4 2 2)	. .	Searching history
(4 3)	.	Organize, create access point
(4 3 1)	. .	Attach category to entity
(4 3 2)	. .	Chunking
(4 3 3)	. .	Organize chronologically
(4 3 4)	. .	Put entity into category
(4 3 4 1)	. . .	File into electronic folders
(4 3 4 1 1)	. . . .	Bookmarking
(4 3 4 2)	. . .	File into paper folders
(4 3 5)	. .	Sorting

(5)		Search history use
(5 1)	.	Search history use by user
(5 1 1)	. .	Learn about system
(5 1 2)	. .	Task management
(5 1 3)	. .	Search history as memory aid
(5 1 3 1)	. . .	Recognition
(5 1 3 2)	. . .	Recreate context
(5 1 3 3)	. . .	Search history as aid to attention
(5 1 3 4)	. . .	Search history as aid to encoding
(5 1 3 4 1)	. . . .	Presentation in structured context
(5 1 3 4 2)	. . . .	Providing clipboard
(5 1 3 5)	. . .	Search history as aid to retrieval
(5 1 3 5 1)	. . . .	Cued recall
(5 1 3 5 2)	. . . .	Recognition vs recall
(5 1 4)	. .	Search history for interpretation
(5 1 4 1)	. . .	Browsing in context
(5 1 4 2)	. . .	Build mental model of problem

(5 1 4 2 1) . . . . . Roadmap of problem
(5 1 4 2 1 1) . . . . . . Interaction of RM w search organizat
(5 1 4 3) . . . . Develop a scheme of legal issues
(5 1 4 4) . . . . Integration of sources
(5 1 4 5) . . . . On-the-fly interpretation
(5 1 4 5 1) . . . . . Review search history
(5 1 4 5 2) . . . . . Review user notes
(5 1 4 6) . . . . Remember tangents and reason for it
(5 1 4 7) . . . . Same legal issue across diff cases
(5 1 4 8) . . . . Search h to reduce mental complexity
(5 1 4 8 1) . . . . . Chunking of information
(5 1 4 8 2) . . . . . External extension of working memo
(5 1 4 9) . . . . Topical exploration
(5 1 4 10) . . . . Writing report
(5 1 4 10 1) . . . . . Integrating materials into documents
(5 1 4 11) . . . Use as evidence
(5 1 5) . . . Using SH to find information
(5 1 5 1) . . . . Integrate results fr multpl searches
(5 1 5 1 1) . . . . . Identify core set of items fr MS
(5 1 5 2) . . . . Meta-search processes
(5 1 5 2 1) . . . . . Knowing when to stop
(5 1 5 2 2) . . . . . Plan and evaluate search
(5 1 5 2 3) . . . . . Reduce disorientation
(5 1 5 2 4) . . . . . Ustand current steps based on SH
(5 1 5 3) . . . . Reuse
(5 1 5 3 1) . . . . . Correct error
(5 1 5 3 2) . . . . . Redo with modification
(5 1 5 3 3) . . . . . Replicate research
(5 1 5 3 4) . . . . . Reuse keywords
(5 1 5 3 5) . . . . . Reuse query
(5 1 5 3 5 1) . . . . . . Modify query
(5 1 5 3 5 2) . . . . . . Rerun same query
(5 1 5 3 6) . . . . . Reuse results
(5 1 5 3 6 1) . . . . . . Reuse document
(5 1 5 3 6 1 1) . . . . Reuse section of document
(5 1 5 3 6 2) . . . . . . Reuse result list
(5 1 5 3 7) . . . . . Reuse URL
(5 1 5 4) . . . . Revisit
(5 1 5 4 1) . . . . . Keeping one's place
(5 1 5 4 2) . . . . . Navigate actions
(5 1 5 4 2 1) . . . . . . Check recent actions

(5 1 5 4 2 2)	. . . . . Going back to previous place
(5 1 5 4 3)	. . . . Recreate context
(5 1 5 4 5)	. . . . Relate current actions to previous
(5 1 5 4 6)	. . . . Review search
(5 1 5 4 6 1)	. . . . . Roadmap of search steps
(5 1 5 4 7)	. . . . Undo-redo
(5 1 6)	. . Administrative use
(5 1 6 1)	. . . Billing
(5 1 6 2)	. . . Defense
(5 1 6 2 1)	. . . . Malpractice suits
(5 1 6 2 2)	. . . . Team accountability
(5 1 6 3)	. . . Time management
(5 1 7)	. . Collaboration
(5 1 7 1)	. . . Defense
(5 1 7 2)	. . . Organizational memory
(5 1 7 3)	. . . Sharing mental models
(5 1 7 4)	. . . Statistics
(5 1 7 5)	. . . Task delegation
(5 1 7 6)	. . . Teamwork
(5 1 7 7)	. . . Training
(5 2)	. Search history use by system
(5 2 1)	. . Automated relevance feedback
(5 2 1 1)	. . . RF Modifying current query
(5 2 1 2)	. . . RF Modifying user profile
(5 2 2)	. . Compare result sets
(5 2 2 1)	. . . Filter out previous results
(5 2 3)	. . Help
(5 2 3 1)	. . . Recognize disfunctional patterns
(5 2 4)	. . Learning for automatic indexing
(5 2 5)	. . Show repetition
(5 2 5 1)	. . . Duplication
(5 2 5 2)	. . . Overlap
(5 2 5 3)	. . . Similarity
(5 2 6)	. . System notes creation
(5 3)	. Recency of history used
(5 4)	. Search history non-use
(6)	Design features
(6 1)	. Control of history recording, mgmt

(6 1 1)	. . Automatic recording
(6 1 1 1)	. . . User annotation on automatic data
(6 1 2)	. . User controlled recording
(6 1 2 1)	. . . Solicited user decision
(6 1 2 2)	. . . System data recorded
(6 1 2 3)	. . . Unsolicited user decision
(6 1 2 4)	. . . User poduced history data
(6 2)	. Data presentations tools, techniques
(6 2 1)	. . Collapse and expand
(6 2 2)	. . Scrolling
(6 2 2 1)	. . . Speed of scrolling
(6 2 3)	. . Timeline
(6 2 4)	. . Visualization
(6 2 5)	. . Zooming
(6 2 6)	. . . Granularity
(6 2 7)	. . . Speed
(6 3)	. Format
(6 3 1)	. . Forms
(6 3 2)	. . Graphical marks
(6 3 2 1)	. . . Type of graphical marking
(6 3 2 1 1)	. . . . Arrow
(6 3 2 1 2)	. . . . Circle
(6 3 2 1 3)	. . . . Cross out
(6 3 2 1 4)	. . . . Draw
(6 3 2 1 5)	. . . . Highlight
(6 3 2 1 6)	. . . . Color
(6 3 2 1 7)	. . . . Underline
(6 3 3)	. . Icons
(6 3 4)	. . Image
(6 3 5)	. . Numeric
(6 3 6)	. . Sound
(6 3 7)	. . Text
(6 4)	. History data presentation
(6 4 1)	. . Granularity of data display
(6 4 2)	. . Separate vs integrated presentation
(6 4 2 1)	. . . History data integrated with item
(6 4 2 2)	. . . Separate history presentation
(6 4 2 2 1)	. . . . History model, order of items
(6 4 2 2 2)	. . . . Search process presentation
(6 4 3)	. . Visual display of query and result
(6 5)	. Interface issues

(6 5 1)	. . Cost benefit issues of interfaces
(6 5 1 1)	. . Customization
(6 5 1 1 1)	. . . Reorganize display
(6 5 2)	. . Flexible manipulation
(6 5 2 1)	. . Preserve context
(6 5 2 2)	. . Privacy protection
(6 5 2 3)	. . Richness of interaction
(6 5 2 3 1)	. . . Amount of data displayed
(6 5 2 3 2)	. . . Degree of interactivity of actions
(6 5 3)	. . Scalability of data display
(6 6)	. Interface tools
(6 6 1)	. . Scratchpad tool
(6 6 2)	. . Organized results collection tool
(6 7)	. Medium
(6 7 1)	. . Computer storage
(6 7 1 1)	. . . Email
(6 7 1 2)	. . . Other files
(6 7 1 3)	. . . Word processing files
(6 7 2)	. . Human memory
(6 7 3)	. . Paper, paper files
(6 7 3 1)	. . handwritten on paper
(6 7 3 2)	. . printouts
(6 8)	. Method of data entry
(6 8 1)	. . Manual data entry
(6 8 1 1)	. . . Keyboard
(6 8 1 2)	. . . Other data entry interface tools
(6 8 1 2 1)	. . . . Button
(6 8 1 2 2)	. . . . Checkbox
(6 8 1 2 3)	. . . . Drawing
(6 8 1 2 4)	. . . . Graphical marking
(6 8 1 2 5)	. . . . Handwriting
(6 8 1 2 6)	. . . . Radio button
(6 8 1 3)	. . . Pointing device
(6 8 1 3 1)	. . . . Drag and drop
(6 8 1 3 2)	. . . . Highlight and click other place
(6 8 2)	. . Voice data entry
(6 9)	. Method of sharing
(6 10)	. Method of task integration

Appendix B. Test protocol used in attorney observations and interviews.

Introduction

Thank you very much for agreeing to participate in this study.

The goal of the project is to find out about the use of search history information in information seeking. A search history is a record of all the steps you took, things you found, notes and markings you made, things you printed out, so it's a track f everything you did, and all text you ad interacted with, including queries you entered, notes you took, documents you found.

I would like to find out whether you keep track of your searches and results somehow now, record your search histories, if you use this information in any ways, or what tools you can imagine or would like to have to use this information later.

I will ask you to search the Westlaw system as you usually do it. While you search, I will ask you to think aloud and tell me what you are doing, why you are doing it, what your reaction is to what you see.

I would like to also ask you to point out any time when the history of previous searches you did or the history of the searching you have done so far would be helpful in any way, and how.

After searching, I will ask you a series of questions about your searching experience, especially about the use of histories.

I will be taping the session, and recording your search log, but your name or identity will not be linked to the data collected in any way. I will also record the steps of your search for further analysis, and make copies of any notes you may take related to the searches and printouts. The record will be handled confidentially and none of the content of the searches will be revealed at any point.

First, I would like to ask you to read and sign this informed consent form. Let me know if you have any questions related to it. Thank you.

Preliminary inquiry

Now I would like to ask you to fill out this short questionnaire. Thank you.

Searching

I will start the tape recorder now.

[start taping]

Before you start searching, can you tell me about the searches you will do, and how they relate to your work and previous searching you have done?

Thank you.

I would like you to start searching now. While you search, please think aloud, and tell me what you are doing, why you are doing it, what your reaction is to what you see.

I would like to also ask you to point out any time when the history of previous searches you did or the history of the searching you have done so far would be helpful in any way, and how.

[whenever you feel the interface could present other information that would be helpful to you in searching

what is extraneous?]

I would like you to keep in mind, that this is not a test of your searching skills or knowledge in any way, but an investigation into the use of search histories, for this reason I would like to ask you to search as you would normally do it, feel free to print, save, take notes, whatever you feel like.

We can take approximately 45 minutes for searching, but you can take more time if you need it. I will sop you after 45 minutes.

Prompts during searching:

What are you thinking?

What do you think happened?

Interview after searching:

Thank you very much. I have a few questions about these searches.

1. Can you summarize the searching you just did?

What do you remember

about the searching in general?

query statements?

results?

intermediate step?

What would you do next?

2. Please look at this list. The items represent different pieces of information about your search. Imagine that you could record everything you do while you search for information and then have it accessible to you later on in a search history display. Please indicate how important you think it would be to record and make available to the searcher this information? Can you please indicate this for the current search session, for longer than a session let's say while you are working on a topic/case, or forever.

Please let me know if you have suggestions for other information.

3. Please look at the list you just went through. If you think about this information and its availability before, during, and after searching, how would you like to use it? What would you want to do with it? Please select ones from the list that you ranked high and describe why you ranked them high and how you would like to use them.

Please select five examples.

4. [Current search history use techniques]

I would like to ask you about your current use of search histories.

4-1. What do you do that helps you remember your searches and results now?

4-2. How do you create and organize reminders? Can you show me some examples?

4-3. Do you use this information again? If yes, how?

[if subject took notes]

If user takes notes ask them about it? Would you like to incorporate that note into the system somehow? How?

[Westlaw, etc. search history mechanisms - if not covered in previous discussion]

Do you use history mechanisms in online searching when those are available in the system? Why? Why not?

Do you ever save your queries? Why? Why not?

Do you save you search results? Why? Why not?

[Westlaw specific]

Do you use the project feature in the Westlaw PC client interface? Why? Why not?

Do you use bookmarks or the 'Go' list in the web interface? Do you use search trails in the new web interface? Why? Why not?

Appendix C. Demographic questionnaire used in attorney observations and interviews.

Age: ______
Gender: ____ **male**
____ **female**
Position: __________________________

QUESTIONNAIRE PART 1: Experience & Information Seeking

1.1 What is your highest level of legal training? ____________________________

1.2 How long have you worked in the legal field? _______ years ______ months

1.3 On the average, how much time do you spend per week searching for information?

__ less than one hour __ 4 to less than 10 hours

__ one to less than 4 hours __ over 10 hours

1.4 How comfortable are you using these search tools?

Boolean operators (AND, OR, NOT)

Very comfortable Not comfortable

1 2 3 4 5 6 7 8 9 NA

West Key Number System

Very comfortable Not comfortable

1 2 3 4 5 6 7 8 9 NA

Natural language searching

Very comfortable Not comfortable

1 2 3 4 5 6 7 8 9 NA

1.5 Have you ever taken a legal research course?

__ Yes. __ No.

1.6 How long have you been using computerized information resources?

__ less than 6 months

__ 6 months to less than 1 year

__ 1 year to less than 2 years

__ 2 years to less than 3 years

__ 3 years to less than 5 years

__ 5 years or more

1.7 How long have you used the Westlaw System? How long ago was this? Can you identify the versions you used?

How long? (yrs or ms)	**How long ago? (yrs)**	**Version***

*Versions:

dedicated terminal	DT
DOS-based Westlaw	DB
Westmate 6.x (PC Client)	W6
Westmate 7.1 (PC Client)	W7
Premise (CD-ROM)	CD
Westlaw.com	WC
Westlaw.com new interface	WN

Which version do you use daily?

1.8 On average, how much time do you spend per week on the Westlaw system?

__ less than one hour __ 4 to less than 10 hours

__ one to less than 4 hours __ over 10 hours

1.9 How many computer applications have you worked with?

(E.g. Microsoft Word, WordPerfect, Netscape, Microsoft Outlook, etc.)

__ none __ 6 - 9

__ 1 - 3 __ 9 - 12

__ 3 - 6 __ more than 12

1.10 Of the following devices, software, and systems, check those that you have personally used and are familiar with:

__ personal computer
__ lap top computer
__ touch screen
__ DVD drive
__ track ball
__ joy stick
__ pen based computing
__ graphics tablet
__ head mounted display
__ modems
__ scanners
__ word processor
__ graphics software
__ spreadsheet software
__ database software
__ computer games
__ voice recognition
__ e-mail
__ Web browser

Appendix D. Sample from Search history item importance questionnaire used in attorney observations and interviews.

Please mark the importance of saving the following history items for the three time periods.

Name of database selected

For the current session

not important very important N/A

1 2 3 4 5

For several sessions/For longer than a session

not important very important N/A

1 2 3 4 5

Forever

not important very important N/A

1 2 3 4 5

Query statement

For the current session

not important very important N/A

1 2 3 4 5

For several sessions/For longer than a session

not important very important N/A

1 2 3 4 5

Forever

not important very important N/A

1 2 3 4 5

Number of results for a query

For the current session

not important very important N/A

1 2 3 4 5

For several sessions/For longer than a session

not important very important N/A

1 2 3 4 5

Forever

not important very important N/A.

Appendix E. Informed consent form used in attorney observations and interviews.

Search Histories for User Support
in Information-Seeking Interfaces

Consent Form

I state that I am over 18 years of age and wish to participate in a program of research being conducted by Anita Komlodi at the College of Library and Information Services, University of Maryland, College Park.

The purpose of the research is to investigate the use of search history information in legal information seeking.

The procedures involve the monitored use of Westlaw. I am willing to perform searches using Westlaw. I am willing to answer open-ended questions about searching and the use of search histories and fill out two questionnaires.

All information collected in the study is confidential, and my name will not be identified at any time.

I understand that I am free to ask questions or to withdraw from participation at any time without penalty.

Anita Komlodi

4105 Hornbake Building

College of Library and Information Services

University of Maryland

College Park, Maryland 20742-4345

(301) 405-0114

____________________________________ ________________

Signature of Participant Date

Appendix F. Sample test protocol used in participatory design sessions.

TRANSCRIPT OF SESSION

Introduction, 10 min-s

- Welcome and thank you very much for agreeing to participate in this study. My name is AK and I am a doctoral student at the College of Information Studies at the University of Maryland. This session is part of my dissertation research on search histories. This is HH/GC, doctoral student from the Computer Science department, who will help me with taking notes, pictures and keeping a score of what we do on the board. He is also interested in the general area of application histories. First of all I would like to give you a short introduction. Please feel free to interrupt me at any time if you have any questions.

- The goal of the project is to find out how searchers use search history information in legal information seeking. A search history is a record of all the steps you took, things you found, notes and markings you made, things you printed out, so it's a track f everything you did, and all text you ad interacted with, including queries you entered, notes you took, documents you found. Imagine that a log of these could be available to you in the computer interface while searching an online or CD ROM information service.

- I would like to find out whether you keep track of your searches and results somehow now, record your search histories, if you use this information in any ways, or what user interface tools you can imagine, or what you think about some of them, or how would like to use this information later.

- Could you please introduce yourselves and tell us about any relevant background you have in the legal or the information field.

- As I said earlier, the goal of this session is twofold. One of them is to find out how you use or could use search history information in our information seeking. The other one is to learn about what kind of software user interfaces could be designed to help users use this search history information that can be automatically recorded.

- As you can see on the agenda, after this introduction we will discuss the concept of search histories in a little more detail. Then we will look at existing production system and research system interfaces that support the use of search history information in information seeking, I will demo these to you and introduce the functionalities and then will ask you to tell me what you think about them, what you like, don't like and how you would use them.

- I will be video and audio taping the session, but your name or identity will not be linked to the data collected in any way. I will also record any designs that you may create. The record will be handled confidentially and none of the content of the searches will be revealed at any point.

- First, I would like to ask you to read and sign this informed consent form. Let me know if you have any questions related to it. Thank you.

- And please fill out this demographic questionnaire so that I can have a picture of your background.

Discussion of search history concept, 10 min-s

- As I mentioned earlier, the goal of my research is to study how legal information seekers, such as attorneys, law librarians, and legal assistants use search history information in their information seeking. By search history information I mean a record of everything or a selection of everything you do while searching. This information can come from just your current sessions, or several previous sessions maybe centered around a topic, or possibly the history of your whole searching life.

- This information can be automatically recorded and provided in the user interface.

- I would like to show you two quick examples of search histories, from the University of Maryland libraries' electronic resources system. The first one is the search history from VictorWEB, the Umd automated library catalog on the web. This is a straight list of all the queries from the search, with number listed before the query showing the number of results returned for the query. If you click on these the system will display that query.

- The second one provides a little more information. The MdUSA system provides access to database from out of campus network computers to members of the campus community. This provides a little more information in a more organized way, or you can also select the result sets and carry out searches within them.

- Have you used the IE history function? This is how it works.

- This log would be recorded by the system and available to the user of the system any time. My main goal for this second part of my dissertation is to come up with designs on how to use and present this search history information in legal information systems.

- Now that you heard about search histories, I have a few questions about whether or how you use search histories.

    Do you use any search histories?

        If yes, how?

        If no, why not?

    Do you any good examples of search history tools?

Demo and critique current search history systems, 30 min-s

- Next I will show you a few different search history interfaces for slightly different applications. I will show you how these work and would like to ask you to tell me what you think about it, how you would use it, what you like about and what you do not like about it.

- I have printed out some questions for you to think about, they are right here, please keep these in mind when looking at the interfaces.

- Please feel free to interrupt and ask any questions or make comments on the features at any point.

WESTLAW HISTORY,

ARIADNE,

PADPRINTS OR OTHER GRAPHICAL WEB HISTORY INTERFACE

TIMELINE PRESENTATION

TOOLTABS FOR SEARCH HISTORIES

What do you think about this tool?

How would you use it?

What additional improvements could be made to make it more useful?

What do you like about it?

What is the best thing about this interface?

What is the worst thing about this interface?

Design of alternative interfaces, 45 min-s

(Explain specific methods to the extent they need to know and show examples.)

Introduce method:

Next I would like to ask you to design search history interfaces that you would like to use in your work. Here is an example of what these interfaces may look like from a similar design session. These do not have to be pretty, but simply show the functionality, there are sticky notes so that you can change the place of different elements. I also made some interface parts (text boxes, buttons, search history interface parts, etc.) that you can stick on you designs. Feel free to experiment and create multiple versions.

You may also take any of the printouts of existing systems and modify those or cut them up and reuse them. Don't forget to include a window/some screen space for the search interface. The goal is to create something similar to the prototype example you just saw.

I would like you to design interfaces that you think would be useful for your own work when searching large legal databases. For example, you can imagine improving on the

Westlaw trail toolset, or just take a Westlaw design screen and create your own search history tools to accompany it.

Also, think about how you would use it to integrate searching with your work, other tasks.

Please explain the interface.

Why did you design it this way?

How would you use it?

Appendix G. Demographic questionnaire used in participatory design session

Age: ______

Gender: ____ male

____ female

Position: ___________________________

1 What is your highest level of legal training? ____________________________

2 How long have you worked in the legal field? ________ years ______ months

3. On the average, how much time do you spend per week searching for information?

__ less than one hour

__ one to less than 4 hours

__ 4 to less than 10 hours

__ over 10 hours

4. How comfortable are you using these search tools?

Boolean operators (AND, OR, NOT)

Very comfortable Not comfortable

1 2 3 4 5 6 7 8 9 NA

West Key Number System

Very comfortable Not comfortable

1 2 3 4 5 6 7 8 9 NA

Natural language searching

Very comfortable Not comfortable

1 2 3 4 5 6 7 8 9 NA

5. On average, how many times a week do you use online or CD ROM information services to find information?

__ Never
__ Once a week
__ 2-3 times a week
__ 5-7 times a week
__ 10 or more times a week

6. On average, how many times a week do you use search history functions? (Save query, save project, bookmarks, browser histories, etc.)

__ Never
__ Once a week
__ 2-3 times a week
__ 5-7 times a week
__ 10 or more times a week

7. Have you ever taken a legal research course?

__ Yes.
__ No.

8. How long have you been using computerized information resources?

__ < 6 months
__ 6 months - 1 year
__ 1 year - 2 years
__ 2 years - 3 years
__ 3 years - 5 years
__ 5 years or more

9. Of the following devices, software, and systems, check those that you have personally used and are familiar with:

__ personal computer	__ track ball	__ head mounted display
__ lap top computer	__ joy stick	
__ touch screen	__ pen based computing	
__ DVD drive	__ graphics tablet	
__ modems	__ database software	
__ scanners	__ computer games	
__ word processor	__ voice recognition	
__ graphics software	__ e-mail	
__ spreadsheet software	__ Web browser	

10 How many computer applications have you worked with? (E.g. Microsoft Word, WordPerfect, Netscape, Microsoft Outlook, etc.)

__ none	__ 6 - 9
__ 1 - 3	__ 9 - 12
__ 3 - 6	__ more than 12

11. When you use a word processor application, like MS Word or WordPerfect, and you
want to copy and paste text, which one is you preference?

__ use the mouse to select copy and paste from the dropdown menu

__ use the mouse to select copy and paste from the icon menu

__ use the CTRL + C, CTRL + V shortcut keys

REFERENCES

Abels, E. G., M. D. White, et al. (1998). "A user-based design process for Web sites." Internet Research 8(1): 39-48.

Allen, B. (1990). "Content analysis and library and information science research." Library and Information Science Research 12: 251-262.

Allen, B. L. (1991). Cognitive research in information science. Annual Review of Information Science and Technology. M. E. Williams. Medford, NJ, Published for the American Society for Information Science by Learned Information, Inc. 26.

Allen, B. L. (1996). Information tasks: Toward a user-centered approach to information systems. San Diego:, Academic Press.

Allen, R. B., P. Obry, et al. (1993). An interface for navigating clustered document sets returned by queries. Conference on Organizational Computing Systems, Milpitas, CA, ACM.

Altmann, E. M. and B. E. John (1999). "Episodic indexing: A model of memory for attention events." Cognitive Science 23(2): 117-156.

Armstrong, C. J. and J. A. Large (1988). Manual of online search strategies. Boston, MA, G. K. Hall & Co.

Ayers, E. and J. Stasko (1995). Using graphic history in browsing the World Wide Web. Proceedings of the 4th International World Wide Web Conference, Boston, MA.

Back, H. B. (1976). The design and evaluation of an interactive reference retrieval system for the management sciences. Business Administration. Pittsburgh, PA, Carnegie-Mellon University.

Barreau, D. K. (1997). Information systems for organization and the problem of ephemeral information. College of Library and Information Services. College Park, MD, University of Maryland at College Park: 191.

Bates, M. J. (1981). Search techniques. Annual Review of Information Science and Technology. M. E. Williams: 139-169.

Bates, M. J. (1989). "The design of browsing and berrypicking techniques for the on-line search interface." Online Review 13(5): 407-431.

Bealieu, M. (1997). "Experiments on interfaces to support query expansion." Journal of Documentation 53(1): 8-19.

Beaulieu, M. and S. Jones (1998). "Interactive searching and interface issues in the Okapi best match probabilistic retrieval system." Interacting with Computers 10(3): 237-248.

Belkin, N., C. Cool, et al. (1995). "Cases, scripts, and information seeking strategies: on the design of interactive information retrieval systems." Expert Systems With Applications 9(3): 379-395.

Belkin, N. J. (1984). "Cognitive models and information transfer." Social Science Information Studies 4: 111-129.

Belkin, N. J. (1990). "The cognitive viewpoint in information science." Journal of Information Science 16: 11-15.

Belkin, N. J., R. N. Oddy, et al. (1982). "ASK for information retrieval: Part I. Background and theory." Journal of Documentation 38(2): 61-71.

Belkin, N. J., R. N. Oddy, et al. (1982). "ASK for information retrieval: Part II. Results of a design study." Journal of Documentation 38(3): 145-64.

Bennett, J. (1972). The user inerface in interactive systems. ARIST. 7.

Berlage, T. (1994). "A selective undo mechanism for graphical user interfaces based on command objects." ACM Transactions on Computer-Human Interaction 1(3): 269-294.

Berring, R. C. (1986). "Full-text databases and legal research: Backing into the future." High Technology Law Journal 1: 27-60.

Best, Catherine P. (2001) Best guide to Canadian legal research. Accessed December 2001. Available from http://legalresearch.org/

Bintliff, B. (1996). "From creativity to computerese: thinking like a lawyer in the computer age." Law Library Journal 88: 338-51.

Blair, D. C. and M. E. Maron (1985). "An evaluation of retrieval effectiveness for a full-text document-retrieval system." Communications of the ACM 28: 289-299.

Borgman, C. L. (1984). Psychological research in human-computer interaction. Annual Review of Information Science and Technology. M. E. Williams, Knowledge Industry Publications, In. 19: 33-64.

Bradley, J. (1993). "Methodological issues and practices in qualitative research." The Library Quarterly 63(4): 431-449.

Card, S. K., J. D. Mackinlay, et al. (1999). Readings in information visualization : using vision to think. San Francisco, CA, Morgan Kaufmann Publishers.

Card, S. K., T. P. Moran, et al. (1983). The psychology of human-computer interaction. Hillsdale, NJ, Lawrence Erlbaum Associates.

Cherry, A. M. (1991). "A measure of CALR (computer-assisted legal research use by first-year law students following mandatory training." Law Library Journal 83: 73-84.

Chiueh, T.-c., T. Mitra, et al. (1998). Zodiac: A history-based interactive video authoring system. ACM Multimedia '98, Bristol, UK.

Cockburn, A. and S. Greenberg (1999). Beyond the 'Back' Button: Issues of Page Representation and Organisation in Graphical Web Navigation Tools. Calgary, Alberta, Canada., Department of Computer Science, University of Calgary,.

Cockburn, A. and S. Jones (1997). Design issues for World Wide Web navigation visualisation tools. Proceedings of RIAO'97: The Fifth Conference on Computer-Assisted Research of Information, Montreal, Quebec, Canada, McGill

University.

Cole, C. and C. Kuhlthau (2000). "Information and information seeking of novice expert lawyers: how experts add value." The New Review of Information Behaviour Research 1.

Collins, S. (1998). "Publishing mergers and controlling costs." The Law Librarian 29(1): 41-42.

Combs, T. T. A. and B. B. Bederson (1999). Does Zooming Improve Image Browsing? Digital Libraries '99, Berkeley, CA, ACM.

Connell, J. and L. Shafer (1995). Object-oriented rapid prototyping. Englewood Cliffs, NJ, Yourdon Press.

Costle, D. M. and C. A. Yirka (1991). "The Law Library of Congress as a national law library." Law Library Journal 83: 219-224.

Creswell, J. W. (1994). A qualitative procedure. Research design: Qualitative and quantitative approaches. Thousand Oaks, CA, Sage: 143-170.

Dabney, D. P. (1986). "The curse of Thamus: An analysis of full-text legal document retrieval." Law Library Journal 78: 5-40.

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. Qualitative research in information management. Englewood, CO, Libraries Unlimited.

Dillon, E. J. (1995). "Hypertext: Xanadu for lawyers - or merely a product looking for something to do." Law Technology Journal 4(2)

Egan, M. A., M. Krishnamoorthy, et al. (1998). FCLUST: A visualization tool for fuzzy clustering. SIGSCE, Atlanta, HA, ACM.

Eisenschitz, T. S. and R. L. Walsh (1995). "Lawyers' attitudes to information." The Law Librarian 26(3): 446-450.

Ellis, D. (1989). "A behavioural model for information retrieval system design." Journal of Information Science 15(4-5): 237-247.

Ellis, D. (1993). "Modeling the information-seeking patterns of academic researchers: A grounded theory approach." Library Quarterly 63(4): 469-486.

Etgen, M. and J. Cantor (1999). What does getting WET (Web Event-logging Tool) mean for web usability? Human Factors & the Web, Gaithersburg, MD, NIST.

Fang, X. (2000). "A Hierarchical Search History for Web Searching." International Journal of Human-Computer Interaction 12(1): 73-88.

Feeman, E. and S. Fertig (1995). Lifestreams: Organizing your electronic life. Working Notes of the AAAI Fall Symposium on AI Applications in Knowledge Navigation and Retrieval, Cambridge, MA, AAAI.

Fidel, R. (1984). "The case study method: A case study." Library and Information Science Research 6: 273-288.

Fidel, R. (1985). "Moves in online searching." Online Review 9(1): 61-74.

Fidel, R. (1990). "Online searching styles: A case-study-based model of searching behavior." Journal of the American Society for Information Science 35(4): 211-221.

Fidel, R. (1993). "Qualitative methods in information retrieval research." Library and Information Science Research 15: 219-247.

Foss, C. L. (1987). Learning from errors in AlgebraLand. Palo Alto, CA, Instiute for Research on Learning.

Freeman, R. R. and P. Atherton (1968). Audacious: An experiment with an on-line interactive reference retrieval system using the Universal Decimal Classification as the index language in the field of nuclear science. American Institute of Physics, AIP/UDC-7, New York, NY.

Garg-Janardan, C. and G. Salvendy (1986). "Contribution of cognitive engineering to the effective design and use of information systems." Informaton services and use 6(5-6): 235-252.

George B. Leeman, J. (1986). "A formal approach to undo operations in programming languages." ACM Transactions on Programming Languages and Systems 8(1): 50-87.

Gonzalez , C. (1996). Does animation in user interfaces improve decision making? Computer-Human Interaction, Vancouver, British Columbia, Canada, ACM.

Gottlieb, L. D., Juris (2001). "User Preferences in the Classification of Electronic Bookmarks: Implications for a Shared System." JASIS 52(7): 517-535.

Greenberg, S. and A. Cockburn (1999). Getting Back to Back: Alternate Behaviors for a Web Browser's Back Button. Conference on Human Factors & the Web, Gaithersburg, MD.

Greenberg, S. and A. Cockburn (1999). Beyond the 'Back' Button: Issues of Page Representation and Organisation in Graphical Web Navigation Tools. Calgary, Canada, Department of Computer Science, University of Calgary:

Greenberg, S. and I. H. Witten (1988). How users repeat their actions on computers: Principles for design of history mechanisms. Conference on Human Factors in Computer Systems, ACM.

Greenleaf, G., A. Mowbray, et al. (1997). Representing and Using Legal Knowledge in Integrated Decision Support Systems. DataLex WorkStations 97.

Griffith, C. (1996). "The Thomson/West merger." Information Today 13(8): 1.

Grosz, B. J. and C. L. Sidner (1986). "Attention, intentions, and structure of discourse." Computational Linguistics 12(3): 175-204.

Halvorson, T. R. (2000). Law of the super searchers: The online secrets of top legal researchers. Medford, NJ : CyberAge Books, 360 pp, ISBN: 091096534X

Hanson, S. J., R. E. Kraut, et al. (1983). "Interface design and multivariate analysis of UNIX command use." ACM Transactions on Office Information Systems 2(1): 42-57.

Harter, S. P. (1992). "Psychological relevance and information science." Journal of the American Society for Information Science 43(602-615).

Harter, S. P. and A. R. Peters (1985). "Heuristics for online information retrieval: a typology ad preliminary listing." Online Review 9(5): 407-424.

Hayes, J. R. (1989). The complete problem solver. Hillsdale, NJ, Erlbaum.

Hearst, M. A. (1995). TileBars: Visualization of term distribution information in full text information access. ACM SIGCHI Conference on Human Factors in Computing Systems, Denver, CO, ACM.

Hearst, M. A. (1997). "Interfaces for searching the web." Scientific American 1997(3).

Hearst, M. A. (1999). User interfaces and vizualization. Modern information retrieval. R. Baeza-Yates and B. Ribeiro-Neto, Adison Wesley Longman Publishing Co. Inc.

Hearst, M. A. and C. Karadi (1997). Cat-a-Cone: An interactive interface for specifying searches and viewing retrieval results using large category hierarchy. SIGIR, Philadelphia, PA, ACM.

Hemmje, M. (1995). LyberWorld: A 3D graphical user interface for fulltext retrieval. SIGCHI, Denver, CO, ACM.

Hert, C. A. (1996). "Oser goals on an online public access catalog." Journal of the American Society for Information Science 47(7): 504-518.

Hert, C. A. (1997). Understanding Information Retrieval Interactions: Theoretical and Practical Implications. Greenwich, CT, Ablex Publishing Co.

Hightower, R. R., L. T. Ring, et al. (1998). Graphical multiscale web histories: A study of PadPrints. HyperText 98, Pittsburgh, PA, ACM.

Hill, B. (1997). "Train hard, search easy--training to use the LEXIS legal research service." The Law Librarian 28: 75-76.

Hill, W. C. and J. D. Hollan (1994). "History-enriched digital objects." The Information Society 10(2): 139-45.

Hill, W. C., J. D. Hollan, et al. (1992). Edit wear and read wear. CHI '92, Association for Computing Machinery.

Hix, D. and H. R. Hartson (1993). Developing user interfaces: Ensuring usability through product and process. New York, NY, John Wiley and Sons, Inc.

Houldsworth, A. and T. S. Eisenschitz (1993). "Computerised litigation support in solicitors' offices." The Law Librarian 24: 82-87.

Ingwersen, P. (1996). "Cognitive prespectives of information retrieval interactions: Elements of a cognitive IR theory." Journal of Documentation 52(1): 3-50.

Jackson, P., K. Al-Kofahi, et al. (1998). Information extraction from case law and retrieval of prior cases by partial parsing and query generation. CIKM, Bethesda, MD, ACM.

Jeffries, J. (1991). "The EC information needs of the legal profession: a case study." The Law Librarian 22: 28-31.

Jenkins, J. G. and K. M. Dallenbach (1924). "Obliviscence during sleep and waking." American Journal of Psychology 35: 605-612.

Johnson-Laird, P. N. (1983). Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Cambridge, MA, Harvard University Press.

Jones, P. M. and K. J. Schneider (1996). "Learning environment for magnetic

resonance spectroscopy (LEMRS): Supporting apprenticeship learning in operational environments."
Journal of Educational Multimedia and Hypermedia 5(2): 151-177.

Kantor, P. B. (1993). "The adaptive network library interface: A historical overview and interim report." Library Hi-Tech 11(3): 81-92.

Katter, R. V. and D. B. McCarn AIM-TWX: An experimental on-line bibliographic retrieval system. D. E. Walker: 159-202.

King, G., H. T. Kung, et al. (1994). The Harvard self-enriching digital library facilities (SELF) project. Digital Libraries '94, College Station, TX.

Koenig, M. E. D. (1990). "Linking library users: A culture change in librarianship." American Libraries 21(9): 844-849.

Korfhage, R. R. (1997). Information storage and retrieval, John Wiley and Sons.

Kuhlthau, C. C. (1991). "Inside the search process: Information seeking from the user's perspective." Journal of the American Society for Information Science 42(5): 361-371.

Kuhlthau, C. C. (1993). Seeking meaning: A process approach to library and information services. Norwood, NJ, Ablex Publishing Corporation.

Kuhlthau, C. C. and S. L. Tama (2001). "Information search process of lawyers: A call for 'Just for me' information services." Journal of Documentation 57(1): 25-43.

Kunz, C. L. (1996). The process of legal research. Boston, MA, Little, Brown.

Lewis, C. and J. Rieman (1993). Task-centered user interface design: A practical introduction. Boulder, CO, University of Colorado.

Lin, X., P. Liebscher, et al. (1991). "Graphical representations of electronic search patterns." Journal of the American Society for Information Science 42(7): 469-478.

Loh, B., J. Radinsky, et al. (1997). The Progress Portfolio: Promoting reflective inquiry in complex investigation environments. Computer Suppored Cooperative Learning '97, Toronto, Ontario, Canada.

Q. S. R. Ltd. (1999). QSR NUD*IST Vivo. Melbourne, Australia, Qualitative Solutions & Research Pty. Ltd.

Mackinlay, J. D., R. Rao, et al. (1995). An organic user interface for searching citation links. SIGCHI.

Maltz, D. and K. Ehrlich (1995). Pointing the way: active collaborative filtering. Conference on Human Factors in Computer Systems, ACM.

Maltz, D. A. and K. Erlich (1995). Pointing the way: Active collaborative filtering. Conference on Human Factors in Computer Systems, Denver. CO, ACM.

Marchionini, G. (1992). "Interfaces for end-user information seeking." JASIS 43(2): 156-163.

Marchionini, G. (1995). Information seeking in electronic environments. Cambridge, Cambridge University Press.

Marchionini, G., S. Dwiggins, et al. (1993). "Information seeking in full-text end-user-oriented search systems: The roles of domain and search expertise." Library and Information Science Research 15: 35-69.

Marchionini, G. and B. Shneiderman (1993). Finding facts vs. browsing knowledge in hypertext systems. Sparks of innovation in human-computer interaction. B. Shneiderman. Norwood, NJ, Ablex Publ.

Marcus, R. S., A. R. Benefeld, et al. The user interface for the INTREX retrieval system. D. E. Walker: 159-202.

Marshall, C. C. P., Morgan N.; Golovchinsky, Gene; Schilit, Bill. N. (2001). Designing e-books for legal research. Joint Conference on Digital Libraries (ACM and IEEE-CS), Roanoke, VA, ACM.

Martin, T. H. (1973). The user interface in interactive systems. ARIST. 8.

Masui, T., M. Minakuchi, et al. (1995). Multiple-view approach for smooth information retrieval. UIST, Pittsburgh, PA, ACM.

Matthijssen, L. J. (1995). An intelligent interface for legal databases, ACM.

Maxwell, J. A. (1996). Qualittive research design: An interactive approach. Thousand Oaks, CA, SAGE Publications.

Meadow, C. T., J. Wang, et al. (1995). "A study of user performance and attitudes with information retrieval interfaces." Journal of the American Society for Information Science 46(7): 490-505.

Mellon, C. A. (1990). The theory underlying naturalistic inquiry. Naturalistic inquiry for library science. New York, Greenwood Press: 1-21.

Miles, M. B. and A. M. Huberman (1994). Qualitative data analysis: An expanded sourcebook. Thousand Oaks, CA, SAGE Publications.

Moens, M.-F. and C. Uyttendaele (1997). "Automatic text structuring and categorization as a first step in summarizing legal cases." Information Processing and Management 33(6): 727-737.

Morse, A. L. (1994). A decision tree for end user searching in a large interactive legal database. National Online Meeting, New York, NY, Learned Information.

Morse, E. (1997). Document vizualization. http://www.sis.pitt.edu/~elm2/DocumentVisualization.html.

Nahl, D. and C. Tenopir (1996). "Affetcive and cognitive searching behavior of novice end-users of a full-text database." Journal of the American Society for Information Science 47(4): 276-286.

Neuman, D. (1992). Checklist of situations for which qualitative methods are appropriate.

Nielsen, J. (1990). Hypertext and Hypermedia: The Internet and Beyond. San Diego, CA, Academic Press.

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

Norman, D. A. and D. E. Rumelhart (1975). Explorations in cognition. San Francisco, California, W. H. Freeman and Company.

O'Day, V. L. and R. Jeffries (1993). Orienteering in an information landscape: how information seekers get from here to there. INTERCHI '93, Amsterdam, The Netherlands, IOS Press.

Pagel, S. B. (1989). "The legal bibliography and user needs." Law Library Journal 81: 387-430.

Park, J. K., Jinwoo (2000). "Contextual Navigation Aids for Two World Wide Web Systems." International Journal of Human-Computer Interaction 12(2): 193-217.

Park, T. and K. Chon (1994). "Collaborative indexing over networked information resources by distributed agents." Distributed Systems Engineering 1(6): 362-374.

Patton, M. Q. (1987). Chapter 2: When to use qualitative methods. How to use qualitative methods in evaluating. Thousand Oaks, CA, Sage: 23-43.

Peek, N. (1997). "Representing Law in Partial Information Structures." Artificial Intelligence and Law 5(4): 263-290.

Plaisant, C., B. Milash, et al. (1996). Life Lines: Visualizing personal histories. Conference on Human Factors in Computer Systems, Vancouver, BC, Canada, ACM.

Plaisant, C., A. Rose, et al. (1999). The design of history mechanisms and their use in collaborative educational simulations. College park, MD, Human-Computer Interaction Laboratory, University of Maryland.

Polya, G. (1957). How to solve it. Princeton, NJ, Princeton University Press.

Poynder, R. (1997). "Lexis-Nexis." Business Information Review 14(3): 109-117.

Prakash, A. and M. J. Knister (1992). Undoing actions in collaborative work. CSCW '92. Conference proceedings on omputer-supported cooperative work.

Prakash, A. and M. J. Knister (1994). "A framework for undoing actions in collaborative systems." ACM Transactions on Computer-Human Interaction 1(4): 295-330.

Pritchard-Schoch, T. N. (1993). "Teaching online legal research. Searching LEXIS and WESTLAW." ONLINE 17: 106-108.

Pritchard-Schoch, T. N. (1994). "Teaching online legal research: LEXIS." ONLINE 18: 106-108.

Quintana, Y., M. Kamel, et al. (1993). Formal method for evaluating information retrieval in hypertext systems. SIGDOC, ACM.

Reed, S. (1992). Cognition. Pacific Grove, California, Brooks/Cole Publishing Company.

Ressel, M., D. Nitsche-Ruhland, et al. (1996). An integrating, transformation-oriented approach to concurrency control and undo in group editors. CSCW '96. Proceedings of the ACM 1996 conference on on Computer supported cooperative work.

Robins, D. (1997). Shifts of focus in information retrieval interaction. American Society for Information Science Annual Meeting, ASIS.

Root, R. W. and S. Draper (1983). Questionnaires as a software evaluation tool. Conference on Human Factors in Computer Systems, Boston, MA, ACM.

Roscheisen, M. and C. Morgensen (1995). "Beyond browsing: Shared comments, SOAPs, trails, and online-communities." Computer Networks and ISDN Systems 28(1-2): 137-46.

Saadoun, A., J.-L. Ermine, et al. (1997). "A Knowledge Engineering Framework for

Intelligent Retrieval of Legal Case Studies." Artificial Intelligence and Law 5(3): 179-205.

Sanderson, R. M. (1990). "The continuing saga of professional end-users: Law students search Dialog at the University of Florida." ONLINE 1990(November): 64-69.

Saracevic, T., H. Mokros, et al. (1990). Nature of interaction between users and intermediaries in online searching: A qualitative analysis. American Society for Information Science Annual Meeting.

Savoy, J. (1994). "Searching Information in Legal Hypertext Systems." Artificial Intelligence and Law 2(3): 205-232.

Schwabach, A. (1997). " An analysis of first-year student Westlaw use: why vendor training isn't enough." Legal Reference Services Quarterly 16(2): 7-54.

Shardanand, U. and P. Maes (1995). Social information filtering: Algorithms for automating 'word of mouth'. Conference on Human Factors in Computer Systems, Denver, CO, ACM.

Shaw, D. (1991). The human-computer interface for information retrieval. Annual Review of Information Science and Technology. M. Williams. Medford, NJ, Learned Information, Inc. 26: 155-195.

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

Shneiderman, B. (1999). Supporting creativity with advanced information-abundant user interfaces. Position Paper for National Science Foundation & European Commission Meeting on Human-Computer Interaction Research Agenda, Toulouse, France.

Shneiderman, B., D. Byrd, et al. (1998). "Sorting out searching: A user interface framework for text searches." Communications of the ACM 41(4): 95-98.

Soergel, D. (1985). Organizing information: Principles of data base and retrieval systems. San Diego, CA, Academic Press.

Solso, R. L. (1995). Cognitive psychology. Boston, MA, Allyn and Bacon.

Sorensen, C., D. Macklin, et al. (2001). "Navigating the World Wide Web: bookmark maintenance architectures." Interacting with Computers 13(3): 375-400.

Spenke, M. (1993). From undo to multi-user applications: the demo. Conference on Human Factors in Computer Systems.

Spink, A. (1994). Term relevance feedback and query expansion: Relation to design. International ACM-SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, ACM.

Spink, A. (1995). "Term relevance feedback and mediated database searching: Implications for information retrieval practice and systems design." Information Processing and Management 31(2): 1-11.

Spink, A. and A. Goodrum (1996). "A study of search intermediary working notes: Implications for IR system design." Information Processing and Management 32(6): 681-695.

Spink, A., A. Goodrum, et al. (1998). "Elicitation behavior during mediated information retrieval." Information Processing and Management 34(2/3): 257-273.

Spink, A. and H. Greisdorf (1997). "Users' partial relevance judgements during online searching." Online & CDROM Review 21(5): 271-279.

Spink, A., T. Wilson, et al. (1998). "Modeling users' successive searches in digital environments: A National Science Foundation/British Library funded study." D-Lib Magazine 1998(April): http://www.dlib.org/dlib/april98/04spink.html#Spi97.

Spink, A. T. S., Tefko (1998). "Human-computer interaction in information retrieval: nature and manifestations of feedback." Interacting with Computers 10(3): 249-267.

Stoan, S. K. (1984). Computer searching: a primer for the uninformed scholar. Library lit. 14--the best of 1983, Scarecrow Press: 76-84.

Sutton, S. A. (1991). Managing legal information: a model of institutional memory based on user cognitive maps. Berkeley, CA, University of California at Berkeley. 45: 186-200.

Sutton, S. A. (1994). "The role of attorney mental models of law in case relevance determinations: An exploratory analysis." Journal of the American Society for Information Science 45(3): 186-200.

Tauscher, L. and S. Greenberh (1997). Revisitaion patterns in world wide web navigation. Conference on Human Factors in Computer Systems, Atlanta, GA, ACM.

Taylor, R. S. (1991). Information use environments. Progress in Communication Sciences. B. Dervin. Norwood, NJ, Ablex Publishing Corporation. 10: 217-255.

Terry, K. R. (1992). "Library needs of a patent attorney." Law Library Journal 84: 117-20.

Timbie, M. and D. Coombs (1969). An interactive information retrieval system: Case studies on the use of DIALOG to search the ERIC Document Files. Institute for Communication Research, Stanford, CA.

Tulving, E. (1985). "How many memory systems are there?" American PSychologist 40(4): 385-398.

Twidale, M. and D. Nichols (1998). "Designing interfaces to support collaboration in information retrieval." Interacting with Computers 10(2): 177-193.

Twidale, M. B. and D. M. Nichols (1996). Collaborative browsing and visualization of the search process. Electronic library and visual information research 3. M. Collier and K. Arnold, Aslib: 51-60.

Twidale, M. B., D. M. Nichols, et al. (1995). "Supporting the active learning of collaborative databse browsing techniques." ALT-J: Association for Learning Technology Journal 3(1): 75-79.

Twidale, M. B., D. M. Nichols, et al. (1995). Supporting the collaborative learning during information searching. Computer Support for Collaborative Learning, Bloomington, IN.

Upshall, Michael. (1991). The Hutchinson encyclopedic dictionary. London, United Kingdom. Hutchinson.

Veerasamy, A. and R. Heikes (1997). Effectiveness of a graphical display of retrieval results. SIGIR, Philadelphia, PA, ACM.

Vickery, B. and A. Vickery (1993). "Online search interface design." Journal of Documentation 49(2): 103-187.

Vigil, P. J. (1986). The software interface. Annual Review of Information Science and Technology. M. E. Williams, Published for the American Society for Information Science by Knowledge Industry Publications, Inc. 21: 63-86.

Visser, P., T. Bench-Capon, et al. (1997). "A Method for Conceptualising Legal Domains." Artificial Intelligence and Law 5(3): 207-242.

Westbrook, L. (1994). "Qualitative research methods: A review of major stages, data analysis techniques, and quality controls." Library and Information Science Research 16: 241-254.

Wexelbat, A. (1998). History-rich tools for social navigation. CHI '98. Proceedings of the CHI 98.

Wexelbat, A. and P. Maes (1997). Footprints: Visualizing histories for web browsing. RIAO '97 Computer-Assisted Information Retrieval on the Internet, Montreal, Canada.

White, M. D. (1981). "The dimensions of the reference interview." RQ 20(4): 373-381.

White, M. D. (1982). The reference encounter model. Theories in library and information science. C. Dusha. New York, NY, Bowker: 38-55.

White, M. D. (1985). "Evaluation of the reference interview." RQ Fall(1985): 76-85.

Wilson, T. D. (1997). "Information behavior: An interdisciplinary approach." Information Processing and Management 33: 551-572.

Wu, M. M. (1993). Information interaction dialog: A study of patron elicitation in the information retrieval interaction. New Brunswick, NJ, Rutgers University.

Yuan, W. (1997). "End-user searching behavior in information retrieval: A longitudinal study." Journal of the American Society for Information Science 48(3): 218-234.