Multi-User Search Engine (MUSE):
Supporting Collaborative
Information Seeking and Retrieval

by

RASHMI KRISHNAPPA

A THESIS
Presented to the Faculty of the Graduate School of the
UNIVERSITY OF MISSOURI-ROLLA
In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN INFORMATION SCIENCE & TECHNOLOGY

2005
Approved by

_____________________________   ______________________________
Dr. Madhu Reddy, Advisor                        Dr. Richard Hall

_______________________________
Dr. Sanjeev Agarwal
ABSTRACT

Collaboration plays an important role in the information seeking and retrieval activities within a team setting. However, there is little explicit support of collaboration in most present-day information retrieval tools. The purpose of this research is to examine the impact of designing features in an information retrieval tool that explicitly supports collaboration. The research was done in two phases. First, I designed and developed a collaborative information seeking and retrieval prototype – MUSE (Multi-User Search Engine). Second, I evaluated the prototype using ten two-member teams to explore how well it supported collaboration during the information seeking and retrieval activities.

Through the evaluation, I found that the collaborative features in MUSE especially the chat function played an important role in enhancing the information seeking and retrieval process for the teams. I also identified unexpected challenges (for instance, trust among the team members) that arose as team members used MUSE during these activities. These challenges were both technical and social in nature.

This research highlights the importance of supporting collaboration during information seeking and retrieval activities. It also starts to identify ways that we can better support these activities through technology.
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all those people who helped me with the successful completion of my research. First of all, I would like to express my gratitude to my advisor Dr. Madhu Reddy, without who this research could be considered incomplete. I thank him for giving me a chance to work with him and for his continued support with valuable advice, and encouragement. His subtle guidance with unbelievable patience has made a great impact on me and this research.

I am grateful to Dr. Arlan R. DeKock, Dean of the School of Management and Information Systems, for supporting me by giving a scholarship to complete my research when I was in real need for it.

I also like to take this opportunity to thank my thesis committee members Dr. Richard Hall, Department of Information Science & Technology and Dr. Sanjeev Agarwal, Department of Electrical & Computer Engineer for their help.

Special thanks to my boyfriend Prabhat Nagaraj, who has had to put up with me over the past few years. None of this would have been possible without his continual support and encouragement. Finally, I would like to thank my parents for everything they have done for me until now without whose love and support, I wouldn’t have been what I am today. I owe all my success to them.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
</tbody>
</table>

## SECTION

1. INTRODUCTION ................................................................. 1
   1.1. COLLABORATIVE INFORMATION RETRIEVAL ........ 1
   1.2. RESEARCH APPROACH ........................................... 2
   1.3. THESIS OVERVIEW ............................................. 5

2. BACKGROUND ................................................................. 6
   2.1. INTRODUCTION .................................................... 6
   2.2. INFORMATION SEEKING AND RETRIEVAL ............... 7
   2.3. COLLABORATION ................................................... 8
       2.3.1. Conventional and Digital Libraries ................. 8
       2.3.2. Collaborative Information Retrieval ................ 9
           2.3.2.1. Collaborative Information Retrieval Project .... 11
   2.4. COLLABORATIVE INFORMATION RETRIEVAL SYSTEMS .... 13
       2.4.1. Ariadne Collaborative Browsing Project .......... 14
       2.4.2. Gopher and MOO ........................................ 15
       2.4.3. FoRSIC .................................................. 16
5.1.3. Time Taken to Reach Goal ................................................................. 40

5.2. ROLE OF CHAT ON COLLABORATIVE INFORMATION RETRIEVAL ......................................................... 42

5.2.1. Why Chat? ....................................................................................... 46

5.2.1.1. Consulting .................................................................................. 46

5.2.1.2. Brainstorming ........................................................................... 47

5.2.1.3. Team Cognition ........................................................................ 47

5.2.2. Related vs. Unrelated .................................................................... 48

5.3. SYSTEM USABILITY ........................................................................ 50

5.3.1. Usability Goals ................................................................................ 50

6. DISCUSSION ......................................................................................... 53

6.1. OVERVIEW ........................................................................................ 53

6.2. SUPPORTING TEAM INFORMATION RETRIEVAL ACTIVITIES 53

6.2.1. Role of Chat .................................................................................... 54

6.2.1.1. Sharing the Search Process ....................................................... 55

6.2.2. Level of Acquaintance .................................................................. 56

6.2.2.1. Level of Expertise ..................................................................... 57

6.2.2.2. Trust ......................................................................................... 57

6.2.2.3. Unrelated Chat .......................................................................... 58

6.3. SYSTEM RECOMMENDATIONS ...................................................... 59

7. CONCLUSION ....................................................................................... 62

BIBLIOGRAPHY ...................................................................................... 63

VITA ........................................................................................................... 66
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1 Collaboration Tool</td>
<td>4</td>
</tr>
<tr>
<td>Figure 3.1 Conceptual Model of MUSE</td>
<td>21</td>
</tr>
<tr>
<td>Figure 3.2 Graphical User Interface</td>
<td>23</td>
</tr>
<tr>
<td>Figure 3.3 Search Window</td>
<td>23</td>
</tr>
<tr>
<td>Figure 3.4 Chat Window</td>
<td>24</td>
</tr>
<tr>
<td>Figure 3.5 Share Window</td>
<td>25</td>
</tr>
<tr>
<td>Figure 3.6 Interface and Backend Components</td>
<td>26</td>
</tr>
<tr>
<td>Figure 5.1 Chat log of team 2</td>
<td>38</td>
</tr>
<tr>
<td>Figure 5.2 Chat log of team 4</td>
<td>40</td>
</tr>
<tr>
<td>Figure 5.3 Chat log of Team 5 highlighting the pattern the team followed</td>
<td>44</td>
</tr>
<tr>
<td>Figure 5.4 Related Chatting</td>
<td>50</td>
</tr>
<tr>
<td>Figure 6.1 The Working of Teams</td>
<td>56</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1 Research and Commercial tools</td>
<td>13</td>
</tr>
<tr>
<td>Table 4.1 Demographic Information of Team Members</td>
<td>29</td>
</tr>
<tr>
<td>Table 5.1 Information about teams and performance</td>
<td>35</td>
</tr>
<tr>
<td>Table 5.2 Time Taken</td>
<td>41</td>
</tr>
<tr>
<td>Table 5.3 Groups, Teams, and Patterns</td>
<td>46</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

Modern information-rich environments have created a need for information retrieval systems that provide an easy and efficient means of seeking and managing information and supporting collaboration. In response to this challenge, research focusing on collaborative information seeking and management activities is growing in the recent years [1, 2, 3].

In organizational environments, collecting, managing, and using information efficiently are critical activities of everyday work. Information retrieval activities are generally viewed as single-user tasks in these settings; individual users seek information to satisfy their particular needs [4]. Consequently, traditional information retrieval systems have been designed for single-users. However, teams are becoming a prominent feature of modern organizations and we must develop information retrieval tools to support their collaborative activities.

Information retrieval involves extracting needed information from vast amounts of available information. Individual users face numerous challenges when seeking information to fulfill their particular information needs [5, 6]. These challenges arise because of the ever-increasing amount of available information that in-turn increases the work load on the individual users [6].

1.1 COLLABORATIVE INFORMATION RETRIEVAL

Teamwork is an integral part of any organizational culture. Effective communication and collaboration among the team members is the key to the success of the team. Since, information retrieval plays a prominent role in organizational work
environments, information seeking and retrieval systems need to focus on supporting team member’s ability to effectively collaborate and work together when seeking information. Improving the effectiveness of formal and informal work teams by providing the users with efficient collaborative information tools is a step towards achieving higher organizational productivity.

Collaboration is integral to information seeking and retrieval activities [2, 5] because in these activities information seekers can benefit from working together to seek and retrieve information [5]. Although collaboration is a crucial element in the information seeking process, traditional information retrieval systems provide limited or no support for collaboration [2, 3].

The objective of this research project is to study the effects of providing explicit computerized support for collaboration within information retrieval systems on the information seeking and retrieval process. To understand these effects, I first developed a prototype that supports collaboration between two users when seeking information. I then evaluated the prototype to explore the effects of the system on the users during the information retrieval process.

1.2 RESEARCH APPROACH

My research focused on studying the collaborative nature of the information seeking and retrieval activities. The research attempted to find a solution that integrates features that promote collaboration with a search interface in order to enhance the collaborative information retrieval process.
The initial goal of the research was to build a prototype that explicitly provided support for collaboration during the information seeking and retrieval process. The next stage was to evaluate the prototype to better understand the collaborative nature of information retrieval process and to guide future design and development of these tools.

The research had two distinct phases.

1. The prototype development phase: This phase constituted the building of a prototype that integrated features to support collaboration between users who were searching for information. The product of this phase was a prototype called **MUSE: Multi-User Search Engine**. MUSE supports multiple users’ abilities to communicate and collaborate synchronously during the process of information retrieval in contrast to the single-user search engines that are available. The prototype supports three primary functions (Figure 1.1):
   a. Information retrieval: linking to a database to facilitate information searching
   b. Chat: A communication medium for the information seekers working as a team
   c. Share: A medium to share the search results

Figure 1.1 illustrates the collaboration tool highlighting the three primary functions.
2. The prototype evaluation phase: The evaluation phase involved testing the prototype by having groups of users evaluate the system. This had three stages:

   a. Observing the users use the system in a controlled environment

   b. Interviewing the users to get input for effective evaluation of the system

   c. Analyze the gathered data for accurate assessment of the goals of the research
1.3 THESIS OVERVIEW

The remainder of the thesis reports on a detailed approach taken to support collaboration in an information retrieval system to improve the efficiency of the information seeking and retrieval process.

The thesis is organized as follows:

- Section 2 Background: This section details studies of collaborative information seeking and retrieval. It also discusses different information retrieval tools.
- Section 3 Prototype Design: This section focuses on the detailed development process followed in the design of MUSE.
- Section 4 Evaluation Methodology: This section focuses on the study methodology used to evaluate the prototype and to measure the effects on team performance in the information retrieval process.
- Section 5 Results of Evaluation: This section discusses the results of the evaluation.
- Section 6 Discussion: This section describes the findings of the study.
- Section 7 Conclusion: The conclusion briefly summarizes findings as they relate to the goals.
2. BACKGROUND

2.1 INTRODUCTION

Information seeking and retrieval plays an important role in human life. The invention of computers initiated sophisticated means of storing and retrieving information [7]. However, even with the expansion of the World Wide Web and networking, information seeking and retrieval activities are still primarily viewed as individual activities. Accordingly, most information seeking and retrieval tools are still designed for single users [2].

The last two decades have experienced a rapid growth in the amount of electronically available information. This rapid growth of information resources makes it difficult for individual users to fulfill their particular information needs [3]. Studies have shown that collaboration reduces the difficulties faced by single users by dividing the work load among the team members; this decreases the time taken to extract the needed information from large amounts of available information [3, 8].

The idea of providing explicit computerized support for collaboration among users during online information retrieval has been a focus of research for the last twenty years [3, 8]. Furthermore, research into developing interface technologies that revolutionize human-human and human-machine collaboration in a network-centric environment has been underway for the past decade [2, 3, 8 and 9]. The objective of this section is to highlight research in the fields of collaborative work and information retrieval. By highlighting this research, we can begin to gain some insight into the
opportunities and challenges regarding the design of truly collaborative information retrieval tools.

2.2 INFORMATION SEEKING AND RETRIEVAL

Information retrieval (IR) activities are a prominent aspect of everyday work. Manber [10] explains that the history of IR dates back to at least 2000 BC, when the first Sumerian literary catalogue was constructed. He argues that information retrieval concepts and problems are not new; they have existed for hundreds if not thousands of years. However, modern information technology has started to change how we approach the problems of information retrieval. Current information retrieval tools provide a variety of basic and advanced features that ease the retrieval of information from the available resources. Some of the advanced features are [11]:

1. Advanced search options – more details regarding the search can be specified to filter the results and get the needed results

2. Advanced keyword search – users can choose from the various options related to the keywords, misspelled keywords and so on.

Still, users of these systems face major limitations [3, 12]. For instance, although collaboration is an important aspect of information seeking and retrieval process [3, 8], these information retrieval tools do not support collaboration between users. Observations of user experiences during information retrieval systems reveal a strong desire for collaborative search [13]. These observations also suggest that support for collaboration is rarely available, and then only in a limited fashion.
2.3 COLLABORATION

Modern communities transcend traditional notions of space and identity [10]. Advances in Internet technology enable people to collaborate with other people, by extending human interaction beyond geographical and chronological boundaries [10]. For instance, people working in organizations can work with colleagues who are located in different geographical locations by exchanging information through worldwide networks.

In an organizational setting, work is distributed among many different teams; each team consists of people with complimentary skills who are committed to a common objective and are accountable for achieving this objective [14]. Organizations strive to build effective and efficient teams that contribute to the overall growth of the organization. Although teams are an important aspect of organizational work, most current information retrieval systems support individual information retrieval practices. These systems do not support collaborative information retrieval practices [3, 8].

People, groups and organizations work together to get desired information through networking, cooperation, coordination, coalitions or collaboration [10]. For instance, people seeking information from online search engines communicate with friends and sometimes even with strangers to get needed information [10].

2.3.1 Conventional and Digital Libraries. The electronic library (also called digital library) has emerged to meet the growing amount and needs of information [2, 3]. It provides a new dimension to the way information retrieval systems are designed. The electronic library is a common metaphor used to represent information retrieval systems [2]. Unlike conventional libraries where people are co-located, electronic libraries need to
address the issues of distance and time. These issues are starting to be addressed by the features of network technologies [2, 3, 15, 16, 17].

The electronic libraries are mostly designed with a single user in mind. Based on observations made in conventional libraries [2, 3], there is a significant amount of collaboration between users within the library environment. It is common for people to seek help from each other when searching for information [2, 3]. The process of information searching has many phases to it [2, 3, 23, 24]:

1. “Consult”: to ask other users for suggestions or information that guide the search process;
2. “Wander”: to look around for various sources of information, which would help get to better results;
3. “Brainstorming”: generate various ideas when using the various resources to find information. These resources could include other users, data sources etc. This could lead to more creativity, and also help reach goals faster;
4. “Bibble”: look through other users work on similar platform.

These studies highlight the important role that the human-human interaction plays in information seeking. They also indicate the need for implementing the naturally occurring collaboration in conventional libraries in information-rich electronic libraries.

2.3.2 Collaborative Information Retrieval. The advent of the WWW\textsuperscript{1} in 1994 initiated researchers’ interest in the collaborative aspects of the information seeking and retrieval process. One focus of interest was on how social networks affected people’s

\textsuperscript{1} http://www.w3.org/History.html
ability to find and assimilate information for their work [18]. At the University of Ulm, researchers [19] developed the CoBrow system to support and explore aspects of collaborative browsing. They stated that

“If someone browses for information, there is a high probability that someone else is interested in the same subject at the same time, but people browsing the WWW are unaware of the presence of any fellow browsers.”

For instance, in such situations, people looking for information on the same subject would achieve more if they worked collaboratively than individually [19]. One reason for this is that the recent rapid proliferation of accessible electronic documents has made it tougher for individual users to retrieve needed information. Collaboration among users can play a significant role in reducing the challenges faced by individual users [20]. Cross’ study [20] points to five categories of benefits of collaboration during information retrieval:

1. Solutions (know what and know how): People can turn to other people to get specific information or answers that addresses their particular needs

2. Meta-knowledge (pointers to databases or other people): People can turn to others to learn about location of relevant information

3. Problem reformulation: turn to others for information that would lead them to think differently about their problem and come up with new ideas

4. Validation of plans or solutions: validate their own solutions or plans from others

5. Legitimation from contact with a respected person: Benefit by the virtue of telling others that they discussed it with the other person; citing a respected source as having reviewed the solution increases its credibility
Cross argues that users of information retrieval systems can benefit in a number of ways from collaboration; he also shows how collaboration is an important aspect of the information retrieval process.

2.3.2.1 Collaborative Information Retrieval Project. Over the years, researchers have explored various means of finding the factors that influence information behavior and how this influences the human – computer interaction [8, 9, 21, 22]. The Collaborative Information Retrieval (CIR) project team at the University of Washington is studying how work teams seek and retrieve information in a collaborative fashion. The three goals of the CIR team were [8]:

1. to analyze the structure of collaboration in selected information-intensive, actual work scenarios and to describe the context that will shape collaborative information retrieval patterns;

2. to explore the nature and occurrence of collaborative information retrieval in work settings and to propose technological innovation and organizational change that could support, facilitate, or improve it;

3. to construct a work-centered conceptual framework that includes collaborative information retrieval.

To achieve these goals, the CIR teams conducted two case studies focusing on two teams from two different organizations. The studies aimed to gain a better insight to understanding the need for collaborative information retrieval and how this process differs between teams working on different projects.
The first case study focused on a design team at Microsoft. The team’s goal was to design a Help and Support Center (HSC) for Microsoft. This would enable the end users of the Microsoft products to rely on a unified, useful and usable help portal [21]. The second case study focused on an airplane system design team at Boeing. The information needs of the two teams differed and each had a different measure for success. The Microsoft team could come up with a version of the system, which could be improved in successive versions. However, the Boeing team had more constraints on achieving an accurate system and could not afford any type of failure. Accordingly, team members interacted differently within the two teams.

A work centered conceptual framework called ‘Cognitive Work Analysis’ [3] was used to study the human – information behavior in the job. There are various dimensions to the Cognitive Work Analysis [3, 8, 9, 21, 22]:

1. The “cognitive dimension”: Information Retrieval process varies based on user’s mental cognition. It also varies based on the nature of the users – novice users, experienced uses, new and old members of the team.

2. The specific task and decision:
   a. The nature of the information sources
   b. The nature of the information needed

3. The organization of the team’s work: The responsibilities assigned to each team member.

4. The organizational culture: The superiority or hierarchy of the team members’ in the organizational structure.
Each of these dimensions were equally important in understanding the collaborative information retrieval process. The studies revealed different factors that influence the need for information and the need for collaboration during information retrieval. The CIR project began to provide a deeper insight to understanding the human-information behavior and gives input to developers who are designing computerized information retrieval systems that explicitly support collaboration.

### 2.4 COLLABORATIVE INFORMATION RETRIEVAL SYSTEMS

Researches have begun to develop information retrieval tools that focus on supporting collaborative work activities. A few commercial tools have also been developed to support this work. Table 2.1 illustrates the features of selected research and commercial tools:

<table>
<thead>
<tr>
<th>Features</th>
<th>Research(R) or Commercial(C)</th>
<th>Information searching</th>
<th>View others previous works (searches)</th>
<th>Share search results</th>
<th>Chat</th>
<th>Multiple users view web page the leader is seeing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ariadne</td>
<td>R</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Gopher and MOO</td>
<td>R</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FORSIC</td>
<td>R</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Recent Versions of Netscape</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enlista</td>
<td>C</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wise Wire</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balze</td>
<td>C</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSN messenger (Beta version)</td>
<td>C</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4.1 Ariadne Collaborative Browsing Project. “…Collaboration is an important aspect of searching online information stores that requires explicit computerized support.” [23]

The “Ariadne Collaborative Browsing” project focuses on providing computerized support for collaborative browsing, information retrieval, and sharing. The researchers’ observations describe how collaboration is an important aspect of information retrieval [2, 23]. For instance, there are three important aspects to the collaborative information retrieval process: (1) performing a search, (2) sharing the search results, and (3) sharing the search process. In an office environment, employees search for information making use of the various sources available. The search process may also include seeking help from colleagues and sharing views and ideas among team members. Collaboration is inherent to an office environment. Yet, as the authors point out, these activities are not well supported by existing computerized information retrieval tools.

The Ariadne system is a tool that was built based on observations made in conventional and electronic libraries [2, 3, 9, 23, 24]. These observations highlighted a significant need for providing computerized support for collaboration during information retrieval. The system provides features for saving and sharing the search process and visualization of the search. This is achieved by capturing the user’s input and recording the database’s output to form command – output pairs.

“These pairs are used to dynamically create records of the search processes, which can be presented at a later time. The system supports synchronous and collocated collaboration facilities, and has been evaluated with subjects from various fields.” [2]
The following observations were made during the evaluation of the system [2, 23]:

1. The search process adopted by users is often ineffective leading to goal failure.

2. The sharing of both the search *product* and the search *process* are important for collaborative activities.

3. Educating the users about using effective search techniques would produce better results.

4. Re-use of previous searches made by other users would increase search effectiveness.

5. Browsing includes searching databases for information as well as for other people who could be a valuable source of information.

6. Users of databases need support to reduce their cognitive load during the search process. This support could be provided externally or internally by the system.

   The Ariadne studies indicate that the users often fail to get the needed information because the search techniques used by the users are ineffective most of the time. The conventional information retrieval systems provide limited support for leading the users in the right track to their goal. The system highlights the significance of supporting collaboration by allowing users to share views and knowledge with each other during the search process.

   **2.4.2 Gopher and MOO.** The merging of Gopher and MOO is another good example of a system that supports collaboration by combining computer-mediated conferencing and online information retrieval [15]. Gopher is an information retrieval tool and MOO provides a social networking environment. At the Xerox Palo Alto
Research Center, Larry Masinter’s work on supporting collaboration within information retrieval systems resulted in a system that combined two different software tools – Gopher and MOO [15]. Single-user retrieval systems lack the support for a link between the social and technical aspects of information retrieval thus creating a gap. This system aims at providing an information retrieval interface to a social networking system, bridging the gap created by single-user information retrieval systems.

2.4.3 FoRSIC. FoRSIC [25] is a research device that attempts to address the issue of information overload by devising a dynamic means of supporting connections between information seekers, information trainers, information tools, and information resources. The research team looked at neglected social factors such as communication and collaboration in the dynamic information retrieval systems. In the field of information science, ontology is a hierarchical structuring of knowledge about things by subcategorizing them according to their essential (or at least relevant and/or cognitive) qualities [25]. The FoRSIC research team describes the role of ontologies in necessitating collaboration in information retrieval systems. The results of evaluating the device with a group of trainers showed that collaboration cannot be forcibly imposed on all users and the users need to be convinced:

“1. to co-specify the constraints of information systems and users behaviors around common languages (ontologies);

2. to take into account (and as equal parts) fine granularity of practices and documents;

3. to be in an explicit frame of knowledge sharing.” [25]
2.4.4 Commercial Tools. Although providing explicit computerized support for collaborative information retrieval is gaining more interest and attention from a research perspective, there are no fully commercialized tools that support this feature. However, commercial software developers have implemented some of the features mentioned above in their systems.

1. Recent versions of Netscape\(^2\) provide a feature that allows multiple users to view the same web page as the leader of the team is looking at.

2. Enlista\(^3\): facilitates users to browse and share information while chatting. Allows user to create simple databases and share them instantly over the internet.

3. WiseWire\(^4\): features that WiseWire provides: searching for information, provide ability to do research, store user actions; users will not have to look through same information more than once, blends innovative search strategies that can be helpful in searching, such as "adaptive collaboration," users will be able to look at what others are looking at.

4. Blaze\(^5\): provides an organizer to help keep track of bookmarks and search results by different users which can be used for further referenced.

5. The new MSN Messenger 7.0 (beta version)\(^6\) is a more enhanced step towards supporting collaboration during information retrieval. The features that this

---

\(^2\) [http://www.netscape.com](http://www.netscape.com)
\(^3\) [http://www.enlista.com/products.html](http://www.enlista.com/products.html)
\(^4\) [http://www.wisewire.com](http://www.wisewire.com)
\(^5\) [http://www.speeditup.com](http://www.speeditup.com)
version provides alongside with the other common messenger features are as follows:

a. Search the web within the same chat window, without having to leave the session

b. The search results can be shared with the other users on the chat session

However, these commercial tools support collaborative information retrieval only to a limited extent. Information-rich environments necessitate the need for more sophisticated and adaptive information retrieval tools. Users working in teams face barriers of time and distance that influence effective communication and collaboration. Although researchers are working on problems in the collaborative information retrieval field, there is still much work to be done in this area.

The major limitation that was identified through the literature review was the lack of support for real-time communication between the users. One of the goals of this project is to design a prototype that provides supports for real-time, synchronous communication between the team members in order to study its impact on information seeking and retrieval process.
3. PROTOTYPE DESIGN

3.1 INTRODUCTION

I developed a prototype of a Multi-User Search Engine (MUSE) to evaluate and explore the efficacy of explicitly supporting collaboration in an information seeking and retrieval system. The design was intended to support the collaborative information seeking and retrieval activities. In this section I will discuss the design process of the prototype MUSE.

3.2 PROBLEM SPACE

Information seeking and retrieval activities are traditionally viewed as single-user tasks. However, in modern information-rich environments, retrieving the needed information is becoming a more complex task for the single-user. Therefore, I wanted to develop and evaluate an information retrieval tool that supports collaborative during the information retrieval process. The design goals were to:

1. Support information seeking and retrieval from a selected database;

2. support real-time, synchronous, text-based communication between two users;

3. Support sharing retrieved results between two users.

Features were designed to support goals (2) and (3) in order to promote collaboration among the team members.
3.2.1 Design Assumptions. There were four assumptions made prior to the design of MUSE:

1. Collaboration is an important aspect of information seeking and retrieval activities.

2. Explicit support for collaboration within information retrieval applications would help users during the information seeking process.

3. Collaboration enhances the information seeking and retrieval activities by reducing the work load on individual users and the time taken to get results.

4. Users would be interested in searching for information while talking to a team member, sharing ideas, sharing search results, consulting each other, and would appreciate a system that would provide all these facilities within one application.

These assumptions laid the foundation for the next stage i.e. to develop a conceptual model of the prototype.

3.3 Conceptual Design

Figure 3.1 shows the conceptual model of the prototype. The order in which the users decide to use the facilities provided by the prototype is dependent on the individual user. As seen in the figure, the users have two entry points. They can either start chatting with the other team member or start searching for information. The users are provided with five options that they can choose to use during their use of the system: (1) chat with each other; (2) search for information (3) save the results of interest for future use (4) retrieve the saved results and (5) share the results. The search function provides support for
retrieving information. The save and retrieve functions are provided to support additional facilities for users to look back at their work. The chat and the share functions are provided in an attempt to help establish and promote collaboration.

3.4 PHYSICAL DESIGN

MUSE is a sophisticated and fully functional high fidelity prototype. MUSE was developed using JAVA which facilitates its integration to the World Wide Web and makes it platform independent. Java Swing components were used for designing the interface. The apache server was used to enable networking between different computers. The code was distributed in 7 files with approximately 900 lines of code. The prototype was developed in two phases:
1. The first phase was to build a link to a selected database that users can search. MUSE is linked to the public medical database ‘PUBMED’. PUBMED is a medical search engine provided by the ‘National Library of Medicine’. It includes over 15 million citations for biomedical articles dating back to the 1950's. PubMed also provides links to other websites that provide full text articles and other related resources. PUBMED is used as a sample database for evaluation the effectiveness of collaboration in information retrieval process. However, MUSE can be linked to any other search engine according to needs.

2. The second phase involved designing a chat application and a share application to support communication between the two users and sharing of the search results in an attempt to promote collaboration among the team members.

3.4.1 Prototype View. The software view of the prototype constitutes of two main components:

1. Interface

2. Backend

3.4.1.1 Interface. The prototype includes a graphical user interface (GUI). Figure 3.2 shows MUSE’s interface. It highlights the three distinct features of the prototype: (1) search (2) chat and (3) share.

1. **Search**: Users type in a keyword to search for information. The search engine retrieves the available results from the database and displays the first twenty results that are retrieved. Figure 3.3 shows the search window.
Figure 3.3 also shows four other buttons within the search window:

1. **Save:** Saves the selected results. The users can select results of their interest using the checkboxes.

2. **Retrieve:** Retrieves the saved results.

3. **Select All:** Selects all the checkboxes.

4. **Clear All:** Deselects all the checkboxes.

2. **Chat:** Users can chat with each other in the chat window. MUSE supports text based messaging between two users. Figure 3.4 shows the chat window of the prototype.

![Figure 3.4 Chat Window](image)

3. **Share:** Team members share the search results with each other in the share window.

The users first select the results they want to share from the retrieved results and then
click the ‘share the search results’ button to share the results. Figure 3.5 shows the share window.

![Share Window Image]

**Figure 3.5 Share Window**

### 3.4.1.2 Backend. The backend of the prototype constitutes of three main components. Figure 3.6 shows the interface and the backend components linked together.

1. **Search Database:** The search function of MUSE is connected to the search database. When the user makes a request by typing in a keyword, the search engine pulls all the available data stored in the database and sends it back to the user. The results are displayed on the screen.

2. **Chat Server:** The chat server is always running in the background waiting for requests. When one of the team members types in a message, the chat server sends it to all the users and the message is displayed on the screens of the team members.
3. **Share Server**: Similar to the chat server, the share server is also running in the background, waiting for requests. Once the user selects the results to be shared, the server forwards it to the other team member and displays it on the screen.

MUSE’s goal is to support collaboration during the information retrieval process. It contains both communication and share features to promote collaboration between two users. We want to allow users to quickly navigate across the interface, chat and share the results with minimal number of button clicks. It is unique compared to other information retrieval tools because of its support for:

1. **Synchronous communication between team members during information retrieval;**
2. Information retrieval, communication and search results sharing within a single interface.

The next three chapters discuss the study methodology involved in evaluating the prototype, followed by understanding the results and their significance.
4. EVALUATION METHODOLOGY

4.1 OVERVIEW

The evaluation of MUSE was done to gain better insight into users’ views and experiences of using a collaborative information retrieval tool. The goals of this evaluation were:

1. examine the advantages and/or disadvantages of collaborating during information retrieval;
2. examine how well MUSE supports collaboration;
3. assess the overall usefulness and the usability of MUSE.

I used two evaluation techniques: observation and interviews to gain insight into the user’s thought process, mental models, and MUSE’s strengths and weakness.

4.2 DATA COLLECTION

The evaluation was performed with ten teams (two users per team). Table 4.1 shows the demographic information of the team members. All the subjects were university students. The teams were asked to perform a set of tasks designed to test the various features of the system. The evaluation of the prototype took each team about 30 to 45 minutes.

---

8 Users were required to sign a study consent form
Table 4.1 Demographic Information of Team Members

<table>
<thead>
<tr>
<th>Team</th>
<th>Team Member</th>
<th>Gender Male-M</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>M</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>F</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>M</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>F</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>F</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>M</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>F</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>F</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>F</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>M</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>A</td>
<td>F</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>M</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>A</td>
<td>F</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>M</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>A</td>
<td>M</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>M</td>
<td>25</td>
</tr>
<tr>
<td>9</td>
<td>A</td>
<td>M</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>F</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>A</td>
<td>M</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>M</td>
<td>18</td>
</tr>
</tbody>
</table>

4.2.1 Search Scenario. The users were asked to use MUSE to achieve the goals of a given scenario. Although no time limit was given, most teams spent between 10 to 30 minutes to achieve the scenario goal. During the scenario, team members were observed using the system. Their use of the system was further analyzed using chat logs and share logs to identify specific communication patterns. Below is the scenario that was given to the users.
You and a friend are chatting over the internet and you listen to the news that there has been an outbreak of a disease.

“Infected adult cattle may develop signs of the disease slowly. It may take from 2 to 8 years from the time an animal becomes infected until it first shows signs of disease. Symptoms in the animal include a change in attitude and behavior, gradual uncoordinated movements, trouble standing and walking, weight loss despite having an appetite, and decreased milk production. Eventually the animal dies. From the onset of symptoms, the animal deteriorates until it either dies or is destroyed (cattle who cannot stand are called “downers”). This disease process may take from 2 weeks to 6 months.”

Your friend asks you more information about it. So you decide to read some papers and share the knowledge with your friend. Earlier during the week when you were surfing news channels on your TV you casually heard an interview by Mr. Sheff (a professor of XYZ University) who was also talking about something similar.

Use the system to find out what these symptoms refer to, an article by Mr. Sheff talking about the disease. Get more information about the same.

The scenario required to team members to work together and arrive at a particular goal. The scenario formed a question as a whole, gave hints on how to progress, and the team members had to find the answer to this question. The goal achieved when the team retrieved information to help them answer the question.

4.3 EVALUATION TECHNIQUES

I used qualitative data collection techniques to gather data. I gathered data through four sources:

1. Observations
2. Interviews
3. Chat logs
4. Share logs
4.3.1 Observations. Teams of two users were observed in a controlled environment. The users were provided with two computers on opposite ends of the room. They were unable to see each other and were given instructions to only communicate with each other using the chat feature of the prototype. The system was set up for the users, and the work process was observed. The users were initially given a brief introduction to the system. The purpose of the study was also explained to the users. During their information seeking process, their communication and searching were recorded in chat and share logs for further analysis. I also recorded the progress of their ability to learn and navigate through the system.

4.3.2 Interviews. The users were interviewed right after using MUSE while things were still fresh in their minds. The interview included structured, semi-structured and unstructured questions. Some questions required yes/no answers, a few required choosing from a set of answers, and others required users’ opinions. The interviews with each of the teams took about 10 to 15 minutes. The interviews were recorded for future analysis. Below are the interviews questions used in the study.

Interview Questions:

1. Have you used a multi-user search engine before?
   
   a. Yes
   
   b. No

   If ‘Yes’ name the software: ______________________

2. In a collaborative search would you prefer to:

   a. chat and then search
b. search and then chat

3. Does chatting with another user influence the search process? Explain?

4. How frequently do you find the need to collaborate during online information retrieval?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Always</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. How frequently did the users share search results?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Always</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. How frequently did the user retrieve and look back at saved results?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Always</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Would you want this system to be extended to include collaboration between more than two users?

8. Did you have problems using the system?

9. Did the system help resolve errors by proving feedback of necessary steps to be taken?

10. Was the interface intuitive?
11. Did you feel any of the buttons used in the system out of place? (Testing
Mapping)

12. Did the tutorial give necessary information to use the system?

13. Suggest improvements to the ‘Help’.

14. Was the system easy to learn? (Testing Learnability)

15. Do you think you would be able to use the system next time without any
guidance? (Testing Memorability)

16. Suggest what other features you would want the system to support.

I used the interview questions to better understand the users’ views about MUSE and their experiences with using the system. The questions also provided a way to confirm the observations I made and get a detailed input from the users.

4.3.3 Chat and Share Logs. The prototype was designed to record the user’s actions for data collection and analysis purposes. I used two types of logged data from each of session of use of the prototype:

1. chat logs: The team chatting was recorded using text files. These chat logs were used during analysis to examine the discussion between users. This was useful in identifying what the team members chatted about, why they chatted, and what they achieved by chatting

2. share logs: All the search results that the team members shared among each other was logged. During the observations, I recorded at what point in time of use of the system the team members shared the search results, and what they achieved by it
The logged data helped in understanding the communication patterns that the teams employed during the scenario.

4.4 DATA ANALYSIS

Large amounts of qualitative data were collected through the evaluations:

1. notes taken during observations and interviews
2. recorded interviews
3. logged files

To analyze the data, I first triangulated the data to understand team members’ views and needs. To triangulate data, I compared the data taken from the different techniques with each other. The goal of triangulation was to ensure the authenticity of the results and to also identify interesting issues for further analysis. Triangulation was also used to identify different patterns in the data. Constant comparison of data from these sources helped me to group data into categories. It also helped me evaluate whether team members actions matched with what they said during interviews. For instance, chat logs showed that in some cases team members chatted about topics unrelated to the search topic. This had a direct impact on the time taken to retrieve the needed information, thus delaying the search process. I was able to compare what I found in the chat logs with the observation I made during their use of MUSE. Triangulation and categorization thus helped during data analysis to sort the data collected and to understand user behavior during collaborative information retrieval. In the next chapter, I will discuss these results of the evaluation based on the analysis of collected data.
5. RESULTS

5.1 TEAMS

The goal of the evaluation was to gauge how well MUSE supported the collaboration during the teams’ information seeking and retrieval activities. The teams were classified based on two different categories (Table 5.1):

1. Team members level of expertise using computerized systems: high, average, low
2. Team members’ level of acquaintance with each other; (1) Users who knew each other and/or worked together in the past; (2) Teams whose users were strangers to each other

<table>
<thead>
<tr>
<th>Team</th>
<th>Team Member</th>
<th>Gender</th>
<th>Age</th>
<th>Level of Expertise</th>
<th>Level of Acquaintance</th>
<th>Time Taken for Task</th>
<th>Search Group</th>
<th>Search Pattern</th>
</tr>
</thead>
</table>
5.1.1 Level of Expertise. Team members’ level of expertise plays a significant role in their view of MUSE. For instance, team member 2b (novice user) stated:

“I am not a computer geek and I don’t like to use more than one system to achieve something, and I don’t like to learn too much. I am old, but information is a basic necessity, and communication is even more important. I am currently working on a project, and I wish I had something like this to use. It does a lot of stuff within one system. I like it this way, simple but still big. Is it coming out soon?”

The users with higher level of expertise had a different opinion about the system. 60% of them were comfortable using the facilities provided by the traditional single-user information retrieval systems, which do not support collaboration. However, the same users also stated that they needed to talk to others, while searching for information, and used other communication tools like messengers, telephones etc to fulfill their communication needs.

Expert User: “It never occurred to me that it would be nice to have an information retrieval system with support for talking to other people, but looking at this system, I feel it is not a bad idea at all.” (Team member 6a)

Expert User: “I will be able to use the facilities provided on the internet to communicate and look for information, but for more novice users, this is an incredible system. When my parents are using the computer, I have to open a web page or something like that and then teach them how to use each of the buttons on the page. For such people, I feel giving more functions on one screen with less work to do but still achieving a lot is a nice idea. The ease with which MUSE allows to chat and look for information is awesome.” (Team member 9a)

Expert User: “I always come across inexperienced users while working on teams. Chatting is a good way to provide help and get help” (Team member 8a)

Novice User: “Information is something that is very important to any individual. When you know there is someone to help you look for it, it saves a lot of time.” (Team member 4b)
The level of expertise of team members had a direct impact on the time taken to retrieve the needed information. The type of communication that they had with each other also depended on the level of expertise of the team members (discussed below).

5.1.2 Level of Acquaintance. All the teams achieved the goal as per the scenario irrespective of the level of acquaintance among the users. However, the path taken to achieve the goal varied between teams constituting of friends and teams of strangers. The teams involving strangers were observed to chat less about unrelated issues and search more. Their chat was mostly related to the subject. More chatting about issues unrelated to the scenario was found in teams in which members knew each other.

Figure 5.1 shows the chat log of team two whose members were strangers to each other. The team took 20 minutes to achieve the scenario goal. Figure 5.1 highlights how the team chatted for approximately 2 minutes and searched individually but frequently shared information for approximately 18 minutes in the span of 20 minutes. The chatting was mostly related to the topic under discussion.

<table>
<thead>
<tr>
<th>Team 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level of acquaintance:</strong> strangers</td>
</tr>
<tr>
<td><strong>Time taken to reach goal</strong> = 20 minutes</td>
</tr>
<tr>
<td><strong>Chat log</strong></td>
</tr>
<tr>
<td>Started chatting…</td>
</tr>
<tr>
<td>Team member 2a : hey</td>
</tr>
<tr>
<td>Team member 2b : yep</td>
</tr>
<tr>
<td>Team member 2b : I have an infected adult cattle. Do you know anything about this?</td>
</tr>
<tr>
<td>Long gap before 2a replies…</td>
</tr>
<tr>
<td><strong>In the meanwhile, both the team members are searching for information…</strong></td>
</tr>
<tr>
<td>Team member 2a : i found an arcile by Dr. Sheff</td>
</tr>
<tr>
<td>Team member 2b : where did you find the article</td>
</tr>
<tr>
<td>Team member 2a : i did da search for &quot;cow disease sheff&quot;</td>
</tr>
<tr>
<td>Team member 2a : in reference to the TV interview</td>
</tr>
</tbody>
</table>
Team member 2a : can you see the shared results?
Team member 2b : yes I can
Team member 2b : how do you actually look at the article
Team member 2a : i don't know
Team member 2a : try to share something

Another long gap before 2b replies...
Both the team members are searching for information...

Finally 2b shares what he found and thinks is the right result
Team member 2b: ok, I think it is the one I put up last
End of session

Figure 5.1 Chat log of team 2

Figure 5.2 shows the chat log of team 4 whose members had worked with each other in the past. The team also took approximately 20 minutes to achieve the scenario goal. The team members chatted for about 10 minutes and searched individually for approximately 10 minutes. The chatting was related to the topic under discussion, but the discussion was more elaborate when compared to the chat session of team 2 (figure 5.1).
Team member 4b: wat disease is that?
Team member 4a: i saw in tv.
Team member 4b: i ve heard abt it and im tryin to get some info
Team member 4b: ooh really?
Team member 4a: spme prof Dr sheff was telling
Team member 4a: it effects the cattle .i am kind of worried.

4b has already started searching while 4a is explaining

Team member 4b: yeah im trying to find some info abt it
Team member 4a: lets search more about it.

Searching and chatting go on simultaneously until they get to the results ...

Team member 4a: i have my farm to think about.
Team member 4b: do u have any clue of what the name is?
Team member 4a: no i dont know the name.
Team member 4b: ok ok forget it
Team member 4a: i typed in keyword animal disease.i got like 20 results.
Team member 4b: there are hundreds of animal diseases
Team member 4a: you have any clue to the name of the disease?
Team member 4b: how can u give the keyword like that?
Team member 4b: nope not yet
Team member 4a: its a new disease ?
Team member 4b: i guess so
Team member 4b: probably they discovered it recently
Team member 4b: so im trying to find out information about it by giving the year
Team member 4b: too along with the other keywords
Team member 4a: i got lot of downer but its showing the name of the person downer
Team member 4b: ooh okay
Team member 4a: i got 1 cow disease .
Team member 4b: found that?
Team member 4a: i dont know if it is the same one.
Team member 4b: yeah it is i guess
Team member 4b: lets search for more results
Team member 4a: Team member 4b i found this 1 do you think this is the one?
Team member 4b: give the symptoms
Team member 4b: did u get any results when u gave the symptoms?
Team member 4b: hello hello knock knock
Team member 4b: u there?
Team member 4a: ya
Team member 4b: what did u give in the search?
Team member 4a: i was able to find only 1 result which i think makes sense.
Team member 4b: thats that?
Team member 4b: why dont u try sharing it?
Team member 4a: i shared with you.
Team member 4a: you saw?
Team member 4a: you found anything?
Team member 4a: wish i could search by year or by symptoms
Team member 4b: what search criteria did u give?no
Team member 4b: yeah...
Team member 4b: but this is by some other user right?/
Team member 4b: i mean author
Team member 4a: ya
Team member 4a: and it is 2004
Team member 4b: what about dr.sheff?
The chat logs shown above indicate the difference in chatting between the teams involving members with different level of acquaintance. Although, both the teams took the same time to reach the goal and the chatting was mostly related to the search topic, there was a significant difference in the amount of chatting among the two teams.

5.1.3 Time Taken to Reach Goal. The success of the teams was determined based on how fast and efficiently the team members arrived at the answer to the question given in the scenario. The time taken by the teams to achieve the goal ranged from 10 minutes to 25 minutes (Table 5.2).
Table 5.2 Time Taken

<table>
<thead>
<tr>
<th>Team</th>
<th>Time Taken (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
</tr>
</tbody>
</table>

During the interview of the team that took maximum time (25 minutes) to get an answer to the given question, team member 5b stated that in a real world environment, the time taken would depend upon various factors like the subject under discussion, clear view of what the goal is and so on.

Below is a part of conversation of a team that was recorded from the chat logs:

Team 5
Team member 5a: thanks for helping me yaar, left to myself I would have taken an hour more
Team member 5b: sure, any time

Team 8
Team member 8a: We might hve spent lot more time switching b/w msngers and mails
Team member 8b: I hope it was voice enabled

Team members affirmed that collaboration with team members reduces the time taken to search for information, since working in teams distributes the work load among the team members.
5.2 ROLE OF CHAT ON COLLABORATIVE INFORMATION RETRIEVAL

One of the goals of evaluation was to understand how important chat was in the information retrieval process. We found that chat played an important role in the search process. This section describes the observations that show why team members chatted in the process of information retrieval and also the challenges faced by team members in the process of communication. The following quotes describe some of the various views on communicating during the information seeking process.

“I would need to talk to someone 90% of the times when I am looking for information. I would be lost without any guidance, to be able to communicate in real time, is surely helpful and reduces a lot of work and mental pressure” (Team member 6a)

“I guess, most of the times, messenger is always on, for help, I am lost always, so I do need help from others, it is better than you trying to find something yourself, it takes less time to just ask someone who you think would know an answer to the question or who will know where to find it” (Team member 1b)

“Chatting helps me for a lot of reasons, for example: to check if what I am searching for is correct or if I am on the right track, makes it faster to achieve goals” (Team member 7a)

The analysis of the observations and interviews clearly identified the chat function as an extremely useful aspect of the interface. It was used more than half of the time by 70% of the teams. Team member 1b said,

“Chat leads the search process, and the share function is like an extended hand to the chat function.” This opinion was shared by most of the teams. Another team member stated: “whenever you are chatting, you do a lot better; it gives you other people’s ideas, which helps get to results faster”

Initially the teams were categorized into three groups based on their understanding of the topic under discussion. Furthermore, the teams seemed to follow four different patterns during the search process.
1. **Group 1**: Both team members had a basic idea of the goal that had to be achieved, but the path to the goal was unclear.

Pattern:

a. The teams discussed the matter

b. Searched for relevant information

c. Shared the findings

d. Asked for the other persons opinion / discussed the matter

Figure 5.3 shows the chat log of team 5 highlighting the pattern the team members followed to search for information.

---

**Chat log of team 5**

**Discussion of the scenario**

Team member 5b: hii
Team member 5b: did u read the scenario ?
Team member 5a: Hi, I have some problem ... New disis is coming!
Team member 5b: yeah , i m reading the scenario too
Team member 5b: havent heard abt this before
Team member 5b: have u ?
Team member 5a: Do you want to read about this disis? I don't want to be with animals who died'd
Team member 5b: haha ,
Team member 5b: yeah , probaby we shld search and look up wat this might be

**Search for information**

**Talk about search**

Team member 5b: THAT GIVES SOME DISEASE ABOUT SOME CATTLE
Team member 5a: Can we take a look at this article?
Team member 5b: MY SEARCH TERM WAS "REDUCED MILK PRODUCTION IN CATTLE "
Team member 5b: PROBABLY
Team member 5a: But if you don't know what is this article about you can not figure it out about disis. ef
Team member 5a: I want ot say that the simptoms are very general and we can find a lot of article, but we never would figure out the disis.u
Team member 5b: so did we arrive on something
Team member 5a: What up about new desis and other deadly stuff?

**Shared the results**

Team member 5b: ok i belive
| Team member 5b : its something to do with Mad cow |
| Discuss the shared results |
| Asking other persons opinion |
| Team member 5b : wat du u say |
| Team member 5b : have u heard abt that before ? |
| Team member 5b : c there is a new author |
| Team member 5b : called bloom |
| Team member 5b : probably we can search something called Sheff |
| Team member 5b : did u try sheff ? |
| Team member 5b : in search |
| Team member 5b : "SHEFF " |
| Team member 5b : try it |
| Team member 5a : I tried "downers" and shered with you the article . |

**Shared the results**

| Discuss the results |
| Team member 5b : yep i saw tat |
| Team member 5b : u think the disease is drowners ? |
| Team member 5b : downers |
| Team member 5a : i will try sheff, but what it is? |
| Team member 5b : looks like sheff is an author |
| Team member 5b : who wrote abt Mad cow diseasse |
| Team member 5b : so i think the disease is Madcow |
| Team member 5b : wat du u say ? |
| Team member 5a : I agree with you |

**End of session**

Figure 5.3 Chat log of Team 5 highlighting the pattern the team followed

2. **Group 2:** One of the team members had a basic idea of the goal that had to be achieved, but the path to the goal was unclear. The other team member was completely dependant on his partner to lead him to the goal.

Pattern 1:

a. One team member asked for help

b. The other member looked for information, while the first member made no attempt to search for anything

c. Shared with the other member/ explained the findings

For instance, Team member 10b said,
“I knew I had to find something, but he was doing a good job, so I just decided to lay back and wait for his opinion.”

Pattern 2:

a. One team member asked for help

b. Both members made an attempt to search for information

c. Shared results with each other / discussed the matter

Team member 8b said
“Chatting is a good way to provide help and get help. Discussing things when I need help helps me.”

3. **Group 3**: Both the team members had a vague picture of the goal to be achieved.

Pattern:

a. searched for information together and discussed the matter

For instance Team member 2b said,

“The scenario talked about some kind of animal disease, and the symptoms suggested that it would affect humans. Initially both of us were lost. But then, somehow we found that it had to be some animal that humans came in close contact with, or used its products. Eventually, we found the word cattle somewhere; finally we landed up with the mad cow disease. Discussing matters with others has always had me think in the right direction”

Table 5.3 shows the grouping of teams based on the observations and patterns the teams followed to reach the goal.
Table 5.3 Groups, Teams, and Patterns

<table>
<thead>
<tr>
<th>Group</th>
<th>Team</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1, 3, 6, 7, 9</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2, 5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4, 8</td>
<td>n/a</td>
</tr>
</tbody>
</table>

One common trait observed in all patterns was that the teams always started with a general idea of the goal. As the teams proceeded, the search process narrowed towards the actual goal.

**5.2.1 Why Chat?** The evaluation showed three interesting reasons for chatting with other users while looking for information.

**5.2.1.1 Consulting.** The chat function was used to seek help and suggestions from the other team members. The team users helped each other at various stages; (i) In the beginning before the search process begins; (ii) Seeking help in the process of searching; (iii) Confirming the results towards the end.

Chat logs:

Team 5

Team member 5b: i have some problem ...New disis is coming!
Team member 5a: what about it?

Team 1

Team member 1a: what do you think abt the article i just shared?
Team member 1b: it does not relate to the symptoms given
The Team member 6b stated:
“chatting is great while looking for information as we can express ourselves on what we are searching, and when assistance is needed during searching, the chatting is an added advantage”

5.2.1.2 Brainstorming. Chatting with another person leads to shared problem solving in which all the team members spontaneously contribute ideas by generating a variety of possible solutions.

Chat logs:
Team 1
Team member 1b: search for BSE, it's another name for Mad Cow
Team member 1a: learn something new every day

Brainstorming results in a number of solutions that the team members discuss about and could also lead to creation of new ideas and views. Communication between team members helped them arrive at the optimal solution to their combined goals thus satisfying their individual needs.

5.2.1.3 Team Cognition. When team members think alike and know how the other team-mates think about the task. The team performed better and the members were more satisfied.

“if there are multiple users in a team, we would not waste time in doing the same things again and again, in searching the same thing, what others have done, I feel it is better to listen to someone than doing all the searching ad reading myself” (Team 4a)

“chatting influences everything I do. To get a feedback from other users, if it was more obscure, and more difficult, I would definitely take more advantage of the system” (Team 2a)
Communication between team members supports and enhances team cognition. Team cognition establishes a common ground for understanding the goals thus promoting collaboration.

### 5.2.2 Related vs. Unrelated Chatting

Given the scenario and time constraints, teams chatted about the search topic about 80% of the time. However, in the interviews, the teams stated that they would behave in different ways according to the real-world situation. For instance, if the team members were close acquaintances, the chatting session might have included more personal comments and jokes. The type of chat could have a direct impact on the time taken to achieve the goal.

“We would probably get to the goal a lot faster if I didn’t know the person meaning lots less jokes and personal conversations” (Team member 2a)

There were a few instances where the teams spent more time conversing about unrelated topics and making comments and jokes within the short span of 10 to 25 minutes of observation.

For instance, one of the teams pretended to be doctors. The team members got completely involved in the talk and got diverted from the actual topic under discussion.

Chat log of team 6

Team member 6a: hey, what you doin tom
Team member 6b: I got an operation to attend!
Team member 6a: are u free for dinner
Team member 6b: I am going out with my bf

Later in the interview the same team members said, “If we were more goal oriented or we had a deadline to work on, talking to the other team members would have
surely helped get to goals faster.” (Team member 6a), Table 5.4 shows parts of chat logs of all the teams to provide a little perspective of their conversations during the scenario. The figure highlights the parts of the chat logs during which team members chatted about the topic related to the search.

<table>
<thead>
<tr>
<th>Team</th>
<th>Chat</th>
</tr>
</thead>
</table>
| Team 1 | A : i found a bunch of stuff on “Sheff”  
B : i tired that and didnt get anything  
B : nevermind, i put Mr. Sheff  
A : yeah i got it  
B : hang on i’ll put it on the shared results  
B : the article by Broxmeyer L.  
A : it also stands for breast self-exam too  
B : ok, a lot of the articles on BSE or prion disease doesn’t have pub med abstracts |
| Team 2 | A : Ian I have an infected adult cattle. Do you know anything about this?  
A : where did you find the article  
B : i did da search for "cow disease sheff"  
B : in reference to the TV interview |
| Team 3 | A : I found something that looks like it.  
B : that’s what i thought it was when i read the paper  
A : How did you read it? All I see was just what is shown in the search results. |
| Team 4 | A : you know there is some outbreak of some disease  
B : wats that??  
A : some animal .  
A : will our cattle be effected in india  
B : wat disease is that?  
A : i saw in tv.  
B : i ve heard abt it and im tryin to get some info  
A : spme prof Dr sheff was telling  
A : it effects the cattle .i am kind of worried. |
| Team 5 | A : Do you want to read about this disis? I don't want to be wiht animals who died!  
B : haha ,  
B : yeah , probaby we shld search and look up wat this might be  
B : THAT GIVES SOME DISEASE ABOUT SOME CATTLE  
A : Can we take a look at this article?  
B : MY SEARCH TERM WAS "REDUCED MILK PRODUCTION IN CATTLE " |
| Team 6 | A : do u think u can help me get more info on it?  
B : just gimme some time on it  
A : i wil also try simultaneously  
B : i have got a few  
B : lemme share it |
| Team 7 | A : but do u think tht is it  
A : given the symptoms  
A : Mad Cow Disease  
B : i just sent u some articles  
A : the one tht u share doesnt mention about prof sheff |
5.3 SYSTEM USABILITY

Overall, the users were happy with the features provided by the system. In this section, we will examine the results of the evaluation of MUSE to see if it meets the usability and user experience goals.

Team 10

Team member 10b: Thanks for helping me buddy
Team member 10a: This sys is kool
Team member 10b: yeah, hope I had it for the project I am working for now.
Team member 10a: lot of data collected involved, I am always lost trying to get the right info out.

User 10b included the same in the interview,

“This is an incredible system, it comes in real handy, its really nice once it is in the market, it is a little late for my research though”

5.3.1 Usability Goals

Learnability: Most users thought that the prototype was easily learnable. In the words of a team member 1b, “The simple layout of the screen elements and the buttons telling you what to do, I found it easy to learn.” Team member 3a stated, “now that you ask, I did not
think there was anything to learn to use on the screen. It was all about the goal we had to reach.”

**Memorability:** The users had to choose from a 3 buttons to use the major functions provided by MUSE, and the buttons were observed to be self explanatory. Team member 7b said, “The interface was standard and I dint have to learn anything. I am sure I will be able to use it next time without any external support.”

**Safety:** Users had few problems using the system. However, since MUSE supports computerized support for collaboration which involves communication between two users who are located geographically apart, privacy is a major issue to be considered. The team’s level of trust about the system has a significant effect on the team’s performance. Team member 9b said “While working on any project, I would be happier if I was sure that my conversation with the other team members remains within us before I can use any system.” Currently, the prototype does not deal with privacy issues, but future enhancements of the system will have look into this matter in more detail.

**User experience goals:** 60% of the users were satisfied with the basic facilities provided by the prototype: ability to communicate with other people during the process of information retrieval. The users suggested that the system should include more features to make it more aesthetically pleasing. One of the users said, “I love what the system supports, but it could include more features that messengers include, like emoticons, font color options and so on.”

The results highlighted the five important aspects of supporting collaborative features within information retrieval systems. The results focused on the importance of chatting in the information retrieval process and how and why people communicated
while searching for information. We also saw how the levels of expertise and the levels of acquaintance among the team members had an impact on the performance of the teams. The team members also faced challenges with respect to deviating from the goal due to discussion of non-topic related subjects. In the next section, we will discuss the implications of these results and their importance.
6. DISCUSSION

6.1 OVERVIEW

The evaluation of MUSE helped us better understand the effects of collaboration during the information retrieval process. The results of the evaluation indicated that collaboration has an impact on information seeking and retrieval activities. The discussion that follows is based on the research objectives posed earlier in the thesis.

1. Design and evaluate a collaborative information seeking and retrieval prototype - MUSE.

2. Explore the efficacy of explicitly supporting collaboration within an information seeking and retrieval system.

6.2 SUPPORTING TEAM INFORMATION RETRIEVAL ACTIVITIES

Information seeking and retrieval activities are highly collaborative [1]. When collaboration is not explicitly supported by information retrieval systems, users find other means of collaborating with each other. These means include tools that support synchronous communication such as instant messengers, telephones and tools that support asynchronous communication such as e-mail and file sharing. One of the subjects stated, “If one software does not provide it, some other software will do. But, it is always nice to have it all in one. I search for information using the tools on the internet, and also use messengers, email or sometimes just a telephone call to discuss the topic if needed.” (Team member 3a) MUSE supports synchronous communication and the ability to share the results of the search. This idea of integrating collaboration with information seeking and retrieval activities was noticeably effective from the team perspective. Team
members said that they would have taken more time to retrieve the same results if they
had worked individually. The required time was further reduced because of the support
provided for chatting and sharing during information searching within MUSE. It would
have taken more time to switch between different applications (or screens) to chat, share
and search for information. The users with higher level of experience and expertise with
using computerized systems used chat to save time and get to results faster. The novice
users saw this as an advanced information retrieval system that was simple to use.
Chatting helped the novice users get help from other team members to search for
information as well as to use the system. Both novice users and the expert users made use
of the support provided by the system for collaboration in ways suited for their particular
needs thus enhancing the efficiency of the search process.

6.2.1 Role of Chat. All team members chatted with each other to: (1) seek help
from each other; (2) discuss the search process and findings; (3) team cognition;   (4)
discuss non search related topics. Team members agreed that it would have taken more
time to individually find the needed information.

Chatting integrated very well with the search process. Team members thought that
chatting was naturally inherent to the search process. None of the ten teams had ever used
a search engine that supported communication or other collaborative features. However,
all the teams frequently came across the need to chat with others while searching for
information under various circumstances in the real world. In these circumstances, the
team members said that they relied on other tools that support collaboration to fulfill their
needs. These other tools included file sharing tools, messengers, e-mail, telephone, etc.
For instance, team member 9b said,
“I have never used a system that allowed me to chat with others when looking for information and I have used forums that support topic specific discussion, where there are a lot of people who post there ideas and views, lot of consulting goes on too, but that is again not synchronous. I have made phone calls when needed and email also helps communicate. This is nice to do both chatting on one side and searching on one side.”

As stated earlier, communication during information retrieval has always existed irrespective of the information retrieval systems support for communication. However, it was mostly done in an ad-hoc fashion.

6.2.1.1 Sharing the Search Process. One of the most interesting findings was that the chat function was used for sharing the search process. It was not only the findings that need to be shared but also the search process. The patterns that the different teams followed in the process of retrieving information highlighted the flow of thought process of each of the team member; they communicated their plans through the chat facility. For instance, when one team member searches for a particular keyword, the findings were shared using the share functionality provided by the system. However, the search process would also include the thoughts of the team member when he/she was seeking this information. Chatting supported this by allowing the team members to:

a. discuss what keyword they used to search

b. discuss what they found

c. discuss what they just shared

d. discuss what they thought about the results
Knowing what a team member is thinking allows information seekers working in a team to better understand and manage the information retrieval process. Figure 6.1 shows the information retrieval process of a team. The input to the team is the need for information. The team members search for information from the database which is an iterative process. They share the search results as well as the search process with each other which is also an iterative process until the team retrieves the needed information.

6.2.2 Level of Acquaintance. The level of acquaintance among the team members had a significant influence on the performance of the teams. We identified three issues that varied with the different levels of acquaintance.
6.2.2.1 Level of Expertise. The results indicated that the level of expertise among the team members had a considerable impact on the team performance. The time taken to achieve the scenario goal varied among the teams based on different levels of expertise. For instance, teams consisting of both team members with lower level of expertise took more time to arrive at the scenario goal compared to the teams with members of higher level of expertise. Furthermore, the impact of the level of expertise of the team members differed among the teams with strangers and teams with close acquaintances. Closely acquainted team members had little difficulty in gauging the expertise level of the fellow team members. However, the team members who were strangers had more difficulty in judging the expertise level of the other team member. Chatting with the team members eventually led to a better understanding of the expertise level of the members and helped promote collaboration. Chatting also allowed users with lower level of expertise to consult users with higher level of expertise for help.

6.2.2.2 Trust. Trust becomes an important factor in certain situations, for instance, searching for information for confidential purposes. In the study, there were two important types of trust: (1) social and (2) technical.

Social trust is trust between team members. Team members said that they would work more effectively as a team, if they could trust each other. Distrust between team
members can create a communication problem that can directly impact their collaboration by limiting the communication. Team member 3b stated,

“Why would I want to talk to someone I don’t trust. It would be more trouble than help. If there is a real need I might talk but still wouldn’t feel comfortable discussing matters that I would have if I trusted him/her more”

This level of trust further decreases when communicating when teams are composed of strangers. Social trust is dependant on the individual team member’s interactions and personalities; there is little that a system can do to promote social trust.

Technical trust focuses on the system’s ability to support team members’ privacy. However, communicating in a virtual environment always comes with the risk of privacy loss. Team member 6b stated “I have used computers all my life. The only way to ensure privacy while using computers on a network is to unplug the computer from the network.” Trusting the system to ensure that it safeguards the teams is important for effective collaboration among team members. The technical issues of trust can be addressed by the system to some extent. For instance, one of the common ways of ensuring privacy is by use of secure passwords. The data sent across the network can also be protected by data encryption and decryption techniques. The system’s assurance of the users’ privacy concerns is important to ensure unobstructed collaboration.

6.2.2.3 Unrelated Chat. Unrelated chatting refers to the talk that is unrelated to the search topic or the search process. Unrelated chatting occurred in both teams with strangers and teams with acquaintances.

a. Reasons for unrelated talk: Chatting with friends or colleagues could lead to discussion about personal comments, jokes, and discussion. On the other hand,
teams involving strangers were observed to chat with each other to learn more about each other.

b. Effects of unrelated talk: Discussion about unnecessary and unrelated topics among acquaintances slowed the process of information seeking. However, chatting between strangers was important to better understand the interests of the other team member in the topic and gauge the expertise level of the other team member which eventually led to better collaboration among the team members.

The issues discussed in the last three sections that arise due to the level of acquaintance have a significant influence on the team’s performance. The expertise level of the team members is known in the case of acquaintances; chatting helps gauge the expertise level in the case of strangers. Teams can take more time to achieve the goal due to discussion of non-search related topics, but in the case of strangers this can lead to better understanding of each other. The team members asserted that the teams would be more successful in seeking and retrieving information in comparison to working as individuals, when all the team members are focused on the goal. They believed that the chat helped them towards their goal by allowing them to share their thought process with each other during the search.

6.3 SYSTEM RECOMMENDATIONS

This section discusses the some recommendations and extensions for the system. Users were generally happy with the system but they did provide some recommendations to enhance already existing features.
• MUSE is currently limited to supporting only two users. Support for collaboration among more than two users would be the first step towards expanding the prototype. 80% of the users felt they could handle multiple users and that such expansion would increase their domain of information retrieval for a single search.

• While providing support for collaboration among more than two users, it is necessary to provide the users more options and control over the system to choose the number of users they wish to communicate with. 20% of the users felt that more than two users would make the process more chaotic and hence decrease their efficiency of information retrieval.

• More share options need to be provided once the prototype is expanded to support more than two users. These options would include:

  a. Choice to select if teams want future teams of the system to view the shared results. This would also include providing an option to let the teams view shared results of other teams

  b. Save and retrieve the shared results for future use

• The prototype can be expanded to provide more search options to serve the following:

  a. to choose the number of results presented on each screen

  b. format in which the users want the results to be displayed like:

    i. abstracts
ii. full articles

iii. hyperlinks

- Enhanced audio and video enabled communication would be a good add-on to support collaboration and save the time taken typing. Team members believed that collaboration is inherent to the information retrieval process, and just text based chatting would not be sufficient.

- The chat can also be enhanced by providing more user friendly options like emoticons, editing options like font size color etc. These features would help the teams in expressing themselves more naturally, making the process of searching and collaborating a more enjoyable experience.

The discussion highlights the importance and need for communication and collaboration in information seeking and retrieval activities. Communication can promote better understanding among the team members especially teams with lower levels of acquaintance. MUSE supported the team members ability to share their thought process and search results. This can enhance the information retrieval process by reducing the time taken to retrieve the needed information and the work load on individual users. The study also pointed out that social behavior can impact the team’s performance, but this social behavior cannot be controlled by the system. Issues such as trust between team members and level of expertise of the team members can affect the way in which the system is used. Furthermore, the goal of the system is not to control the user behavior but to support and enhance collaboration between users during the information retrieval process.
7. CONCLUSION

The last decade has seen significant changes in the field of information retrieval with the expansion of the World Wide Web. The volume of information available to individuals in the electronic form is ever-increasing. Despite the growth of the information retrieval field, there has not been a great deal of attention paid to developing information retrieval systems that support collaboration amongst users. My research focused on studying the collaborative nature of information seeking and retrieval activities by (1) building an information retrieval prototype that explicitly supports collaboration and (2) evaluating the system to understand how well it supported collaboration.

The prototype MUSE was built to support team information retrieval activities. MUSE is a search engine that supports information seeking and retrieval activities with additional support for communication through chat and share options. These additional features were integrated within the prototype to promote team collaboration.

The results of evaluation of MUSE point to the importance and need for collaboration during information retrieval. Chatting played a significant role in promoting collaboration among team members. Team members chatted with each other to communicate their thought process and search results; this led to better understanding of both the search process and the findings.

This study found that collaboration reduced the challenges faced by users seeking information by distributing the work load to multiple users. MUSE with its communication and share features helped promote collaboration and support the users’ information retrieval activities.
BIBLIOGRAPHY


Rashmi Krishnappa was born in Bangalore, India, on March 10, 1982. In December 2003, she received her B.S. with Honors in Computer Science and Engineering from the Visveswaraiah Technical University, Bangalore, India. She has worked as a graduate research assistant from August 2004 till June 2005 in the department of Information Science and Technology (IST). In June 22, 2005, she received her M.S. degree in Information Science and Technology from the University of Missouri-Rolla, Rolla, Missouri, USA.