

December 2013

Investigating User Search Tactic Patterns and System Support in Using Digital Libraries

Soohyung Joo

University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>



Part of the [Library and Information Science Commons](#)

Recommended Citation

Joo, Soohyung, "Investigating User Search Tactic Patterns and System Support in Using Digital Libraries" (2013). *Theses and Dissertations*. 360.

<https://dc.uwm.edu/etd/360>

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

**INVESTIGATING USER SEARCH TACTIC PATTERNS AND
SYSTEM SUPPORT IN USING DIGITAL LIBRARIES**

by

Soohyung Joo

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
in Information Studies

at

University of Wisconsin Milwaukee

December 2013

ABSTRACT
INVESTIGATING USER SEARCH TACTIC PATTERNS AND SYSTEM
SUPPORT IN USING DIGITAL LIBRARIES

by

Soohyung Joo

The University of Wisconsin-Milwaukee, 2013
Under the Supervision of Professor Iris Xie

This study aims to investigate users' search tactic application and system support in using digital libraries. A user study was conducted with sixty digital library users. The study was designed to answer three research questions: 1) How do users engage in a search process by applying different types of search tactics while conducting different search tasks?; 2) How does the system support users to apply different types of search tactics?; 3) How do users' search tactic application and system support for different types of search tactics affect search outputs? Sixty student subjects were recruited from different disciplines in a state research university. Multiple methods were employed to collect data, including questionnaires, transaction logs and think-aloud protocols. Subjects were asked to conduct three different types of search tasks, namely, known-item search, specific information search and exploratory search, using Library of Congress Digital Libraries. To explore users' search tactic patterns (RQ1), quantitative analysis was conducted, including descriptive statistics, kernel regression, transition analysis, and clustering analysis. Types of system support were explored by analyzing system features for search tactic application. In addition, users' perceived system support, difficulty, and satisfaction with search tactic application were measured using post-search questionnaires (RQ2).

Finally, the study examined the causal relationships between search process and search outputs (RQ 3) based on multiple regression and structural equation modeling.

This study uncovers unique behavior of users' search tactic application and corresponding system support in the context of digital libraries. First, search tactic selections, changes, and transitions were explored in different task situations – known-item search, specific information search, and exploratory search. Search tactic application patterns differed by task type. In known-item search tasks, users preferred to apply search query creation and following search result evaluation tactics, but less query reformulation or iterative tactic loops were observed. In specific information search tasks, iterative search result evaluation strategies were dominantly used. In exploratory tasks, browsing tactics were frequently selected as well as search result evaluation tactics. Second, this study identified different types of system support for search tactic application. System support, difficulty, and satisfaction were measure in terms of search tactic application focusing on search process. Users perceived relatively high system support for accessing and browsing tactics while less support for query reformulation and item evaluation tactics. Third, the effects of search tactic selections and system support on search outputs were examined based on multiple regression. In known-item searches, frequencies of query creation and accessing forwarding tactics would positively affect search efficiency. In specific information searches, time spent on applying search result evaluation tactics would have a positive impact on success rate. In exploratory searches, browsing tactics turned out to be positively associated with aspectual recall and satisfaction with search results. Based on the findings, the author discussed unique patterns of users' search tactic application as well as system design implications in digital library environments.

© Copyright by Soohyung Joo, 2013

All Rights Reserved

TABLE OF CONTENTS

ABSTRACT	ii
LIST OF TABLES	vii
LIST OF FIGURES.....	ix
ACKNOWLEDGEMENTS	xi
CHAPTER 1. INTRODUCTION	1
1.1. Background	1
1.2. Significance of the Study	5
1.3. Research Questions	8
1.4. Design of the Study	9
1.5. Dissertation Overview	10
1.6. Chapter Summary	10
CHAPTER 2. LITERATURE REVIEW	11
2.1. Definitions of Major Terms	11
2.2. User Engagement and System Support	16
2.3. Search Tactics	19
2.4. Information Retrieval Evaluation	23
2.4.1. System-Oriented IR Evaluation	23
2.4.2. User-oriented IR Evaluation	27
2.5. Interactive IR Evaluation	31
2.5.1. TREC Interactive Track	32
2.5.2. Evaluation Criteria in Interactive IR	34
2.5.3. Search Tasks in Interactive IR	36
2.5.4. Digital Library System Evaluation	38
2.6. Information Seeking Process Models	43
2.7. Limitations of Previous Research	52
2.8. Conceptual Research Model of User-System Interactions in Digital Libraries	54
2.9. Chapter Summary	59
CHAPTER 3. METHODOLOGY	61
3.1. Sample	61
3.2. Data Collection Methods and Procedures	63
3.2.1. Data Collection Methods	63
3.2.2. Selected Digital Library System	65
3.2.3. Search Tasks	66
3.3. Data Analysis	68
3.4. Chapter Summary	82

CHAPTER 4. RESULTS	83
4.1. Users' Search Tactic Application	83
4.1.1. Frequency of Search Tactics Applied in Three Different Tasks	83
4.1.2. Time Spent on Applying Search Tactics	93
4.1.3. Search Tactic Probability Change in a Session	100
4.1.4. Transitions in Search Tactics	105
4.1.5. Clustering Search Sessions by Search Tactic Selection	114
4.2. System Support for User Search Tactics	119
4.2.1. Types of System Support for Search Tactics	119
4.2.2. Measuring User Perceptions of Search Process	121
4.2.3. Measuring Users' Perceived Satisfaction with Search Process	128
4.3. Effects of User Search Tactic Selection and System Support on Search Outputs	135
4.3.1. Effects of Search Tactic Selections on Search Efficiency - Task 1	135
4.3.2. Effects of Search Tactic Selections on Task Success Rate - Task 2	138
4.3.3. Effects of Search Tactic Selections on Aspectual Recall and Satisfaction with Search Results - Task 3	143
4.3.4. Effects of Search Processes on Search Outputs	147
4.4. Chapter Summary	154
CHAPTER 5. DISCUSSION	156
5.1. Users' Unique Search Patterns and IR System Design Implications	156
5.1.1. Effects of Search Task Type on Search Tactic Selections	156
5.1.2. Search Tactic Changes over Time in a Session	160
5.1.3. Frequently Applied Search Strategies	162
5.1.4. Infrequently Applied Search Tactics	164
5.1.5. Effects of Search Process on Search Output	166
5.2. Methodological Implications in interactive IR	168
5.3. Limitations of the Study	170
5.4. Chapter Summary	172
CHAPTER 6. CONCLUSIONS	174
REFERENCES	179
APPENDIX 1	207

LIST OF TABLES

Table 2-1. Definitions of key terms and concepts	15
Table 2-2. The comparison of usability attributes by different researchers	31
Table 2-3. Saracevic's digital library evaluation framework	40
Table 2-4. Faceted classification of information-seeking strategies	47
Table 2-5. Coding scheme of search tactics	57
Table 3-1. Proportions of research subjects by major and academic status	62
Table 3-2. Demographic characteristics of subjects	63
Table 3-3. Three search task types: known-item search, specific information search, and exploratory search	68
Table 3-4. Data analysis methods	69
Table 3-5. System support evaluation criteria and associated measures	81
Table 4-1. Frequency of each type of search tactic in Task 1	85
Table 4-2. Frequency of each type of search tactic in Task 2	87
Table 4-3. Frequency of each type of search tactic in Task 3	89
Table 4-4. Comparison of frequency of each type of search tactic by task type	92
Table 4-5. Time spent on search tactics in Task 1	94
Table 4-6. Time spent on search tactics in Task 2	96
Table 4-7. Time spent on search tactics in Task 3	97
Table 4-8. Comparison of time spent on search tactics by task type	99
Table 4-9. Model fit of Kernel regression estimation (Task 1)	101
Table 4-10. Model fit of Kernel regression estimation (Task 2)	103
Table 4-11. Model fit of Kernel regression estimation (Task 3)	105
Table 4-12. Frequency and probability matrix of search-tactic transitions (Task 1)..	106
Table 4-13. Frequency and probability matrix of search-tactic transitions (Task 2)..	109
Table 4-14. Frequency and probability matrix of search-tactic transitions (Task 3)..	112
Table 4-15. Types of system support for Creat tactics in LOC-DL	120
Table 4-16. Types of system support for EvalR tactics in LOC-DL	122
Table 4-17. Types of system support for Mod tactics in LOC-DL	123
Table 4-18. Types of system support for EvalI tactics in LOC-DL	125
Table 4-19. Types of system support for Xplor tactics in LOC-DL	126

Table 4-20. Types of system support for AccF tactics in LOC-DL	127
Table 4-21. Types of system support for AccB tactics in LOC-DL	128
Table 4-22. Users' perceived system support for each type of search tactic	129
Table 4-23. Mean difference of pairwise comparisons between tactics	130
Table 4-24. Users' perceived difficulty in applying each type of search tactic	131
Table 4-25. Mean difference of pairwise comparisons between tactics	131
Table 4-26. Correlation between system support and difficulty	132
Table 4-27. User satisfaction in applying each type of search tactic	134
Table 4-28. Mean difference of pairwise comparisons between tactics	134
Table 4-29. Correlation between satisfaction, system support, and difficulty for each type of search tactic (Pearson r)	135
Table 4-30. Correlation coefficients between search tactic selection and search efficiency (Pearson r)	137
Table 4-31. Regression of session length on search tactic frequency in Task 1 (Stepwise method)	138
Table 4-32. Correlation coefficients between search tactic selection and success rate (Pearson r)	139
Table 4-33. Regression of success rate on search tactic selection in Task 2 (Stepwise method)	140
Table 4-34. Correlation coefficients between time spent on search tactics and success rate (Pearson r)	141
Table 4-35. Regression of success rate on time spent on search tactics in Task 2 (Stepwise method)	141
Table 4-36. Correlation coefficients between search tactic selection and search outputs (Pearson r)	143
Table 4-37. Regression of aspectual recall on frequencies of EvalI and Xplor tactics (Stepwise method)	144
Table 4-38. Regression of satisfaction with search results on frequency of EvalI tactics (Stepwise method)	144
Table 4-39. Correlation coefficients between time spent on search tactics and search outputs (Pearson r)	145
Table 4-40. Regression of aspectual recall on time spent on search tactics	146
Table 4-41. Regression of time spent on applying search tactics on satisfaction with search results	147
Table 4-42. Correlation coefficients between the perceived system support and search outputs (Pearson r)	148
Table 4-43. Regression of perceived system support on aspectual recall	149
Table 4-44. Regression of satisfaction with search tactics on aspectual recall	150
Table 4-45. Regression weights estimated in the structural path model	153
Table 4-46. Correlations between the exogenous variables	154

LIST OF FIGURES

Figure 2-1. A nested model of information behavior, seeking, interactive information retrieval, and information retrieval	15
Figure 2-2. Research continuum for conceptualizing interactive IR research	31
Figure 2-3. Model of optimal support of ease of use and user control.....	40
Figure 2-4. Marchionini's Information Seeking-Process Model.....	47
Figure 2-5. The basic laboratory model of IR	57
Figure 2-6. DL evaluation framework: Criteria, problem, and implications for design	62
Figure 2-7. A stage process model of Ellis' information behavioral framework	63
Figure 2-8. A berry-picking, evolving search	68
Figure 2-9. Belkin's episode model of interaction with texts: a model of IR support for multiple interactions	69
Figure 2-10. Vakkari's model of the information search process in task performance	81
Figure 2-11. Planned-situational interactive IR model	85
Figure 3-1. Data collection procedures	87
Figure 3-2. Example of kernel regression modeling	89
Figure 3-3. Examining the effect of search tactic selection on session efficiency	92
Figure 3-4. Examining the effect of search tactic selections on success rate	94
Figure 3-5. Examining the effect of search tactic selections on IR outputs	96
Figure 3-6. Examining the effect of system support on IR outputs	97
Figure 4-1. Frequency of each type of search tactic in Task 1	99
Figure 4-2. Frequency of each type of search tactic in Task 2	101
Figure 4-3. Frequency of each type of search tactic in Task 3	103
Figure 4-4. Comparison of frequency of each type of search tactic per minute	105
Figure 4-5. Average time spent on each type of search tactic per session in Task 1..	106
Figure 4-6. Average time spent on each type of search tactic per session in Task 2..	109
Figure 4-7. Average time spent on each type of search tactic per session in Task 3..	112
Figure 4-8. Comparison of time spent on search tactics by task type.....	120
Figure 4-9. Kernel regression of search tactic selection in a session: Task 1.....	122
Figure 4-10. Kernel regression of search tactic selection in a session: Task 2	123

Figure 4-11. Kernel regression of search tactic selection in a session: Task 3	125
Figure 4-12. Search tactic transition model in Task 1	126
Figure 4-13. Search tactic transition model in Task 2	127
Figure 4-14. Two frequently applied search strategies identified from tactic transition analysis: Iterative browsing and iterative search results evaluation	128
Figure 4-15. Clustering analysis of sessions in Task 1	129
Figure 4-16. MDS map of search sessions in Task 1.....	130
Figure 4-17. Clustering analysis of sessions in Task 2	131
Figure 4-18. MDS map of search sessions in Task 2	131
Figure 4-19. Clustering analysis of sessions in Task 3	132
Figure 4-20. MDS map of sessions in Task 3	134
Figure 4-21. Comparison between perceived system support and difficulty levels in applying search tactics	134
Figure 4-22. A structural path model between search process variables and search output variables	135

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Iris Xie, for all of her extraordinary support and guidance from the research planning and development of this dissertation. Without her great intelligence and exceptional guidance, I would never have been able to complete this dissertation. She also provided me with unwavering encouragement and support in various ways. Also, I would like to thank all my dissertation committee members, Dr. Dietmar Wolfram, Dr. Xiangming Mu, Dr. Pertti Vakkari, and Dr. Diane Kelly. I am in debt to all their valuable advice and comments, which definitely helped to improve the quality of this work.

I am also grateful for my SOIS doctoral colleagues who provided support and encouragement to me. They were always willing to help and give their best suggestions. Especially, I would like to thank Chunsheng Huang for developing research ideas and collecting data together from the beginning of this project. I am thankful to Kun Lu for his help and friendship throughout the PhD program.

I wish to thank my parents for their unconditional love. They have always believed in me and supported me reach my goals. Also, I would like to thank my son, Andrew. He has been a constant source of energy and incomparable joy. Finally, and most importantly, I would like to thank my wife, Jaehyun, for her support, encouragement, patience and unwavering love. Without her support and love, I wouldn't be where I am today.

CHAPTER 1. INTRODUCTION

1.1. Background

Information retrieval (IR) by nature is a series of interactions between a user and an information system. In order to obtain relevant information from an information system, users should engage in various types of sub-tasks during a search process, ranging from defining a problem, formulating queries, browsing categories, examining results, evaluating relevance, and to extracting information (e.g., Bates, 1990; Marchionini, 2006; Xie, 2008a). Traditionally, in a narrow sense, IR research has focused on matching queries against a store of indexed texts (Robins, 2000). From the perspective of system side approaches, IR research has studied how precisely and completely text representations match with users' search terms. In real situations, however, user roles are not limited to query creation and consequent search result judgment. Users access information through topic categories provided by the system even without creating any query, and users judge the relevance of individual information items as well as search result pages (Xie & Benoit, 2013). IR tasks can be accomplished by users' different types of search strategies while interacting with the information system.

Researchers began to pay attention to various types of user engagement in the search process and have tried to understand how users interact with the information system in different search tasks. This line of studies has emerged as a distinct research area called interactive IR. While traditional IR studies abstract users out of the IR evaluation model, interactive IR research focuses on users' behaviors and experiences and interactions between users and systems and users and information (Kelly, 2009). Interactive IR has widened the research objects and diversified methodologies in IR research. In addition,

interactive IR incorporates various contextual factors, such as task types, user characteristics, user knowledge, search stages, and system features, in IR studies.

Interactive IR research assumes that understanding users and associated factors should be the fundamental base to design effective IR systems in realistic situations, considering various search tasks and taking into account dynamic user-system interactions (Xie, 2008a).

In spite of the benefits from interactive IR approaches, there is a smaller body of research on interactive IR evaluation models, methods and metrics (Kelly, 2009). Although interactive IR is one of the growing research groups in the field of information science during the last decade, fewer studies have been conducted with regard to quantitative modeling, evaluation and measurement compared to the classic IR research area. In addition, research on interactive IR evaluation is still too exclusively focused on the quality of search results, not search processes (Järvelin, 2009). Although interactive IR has extended the research scope by including users, most interactive IR studies still remain in measuring how search results are relevant, useful or satisfactory as an indicator for system performance (Borlund, 2003).

This dissertation focuses more on search process rather than search outputs. This study intends to quantitatively model users' search processes and assess system support for user interactions in using digital libraries. Also, this study tries to suggest a new interactive IR research method that explores user engagement patterns at the unit of search tactics. Additionally, relationships between search processes and search outputs are further examined, in particular how users' search tactic application behavior would influence search outputs.

The basic assumption of this study is that IR consists of search process and resulting search outputs. The main purpose of this study is to explain how users engage in a search process by applying different types of search tactics, and to assess in what ways and to what extent the system supports users' search tactic application. Search outputs are a marginal research interest of this study. While previous studies mostly tried to examine the effects of specific system feature uses (e.g., comparing two interfaces) or search tasks on search output, this study attempts to examine the influence of search tactic application on search output.

In this study, search tactics are selected as a unit of analysis to investigate search processes. Search tactics are a fundamental granularity of user engagement that enables users to accomplish a specific objective during a search process. Also, search tactics are the unit of user actions that involve a users' certain intention (Xie, 2008a). In information sciences, users' behaviors have been analyzed in different levels, such as search strategies (e.g., building block, pearl growing, information foraging, etc.), search tactics (query creation, browsing, result evaluation, etc.), or search moves (queries, clicks, page views, dwell time, etc.). Search strategies are a higher level concept, so it has been challenging to mathematically trace their patterns in a search session. Search moves have been most widely adopted in interactive IR research because it is relatively easy to measure. The concept of search tactics started to be discussed in late 1970s (Bates, 1979). After that, many researchers identified different types of search tactics (e.i., sometimes called sub-tasks, search actions, or search intention), and employed different types of search tactics in formulating their information seeking models (Marchionini, 1995; Belkin, 1996; Ellis & Haugan, 1997; Vakkari, 2001; Xie, 2008a; etc.). Recently, search tactics have been

frequently adopted in quantitative user modeling studies in interactive IR research (e.g., Kules & Capra 2011; Yue et al. 2012; Han et al. 2013; Fujikawa et al. 2013; Lin & Xie 2013). This study also quantitatively analyzes users' information search sessions at the level of search tactics. Search tactics give more information than search moves while quantitative modeling is still technically feasible.

In this study, the author selected a currently operating information system to be investigated, instead of an experimental system. Digital libraries emerged as a unique information system that extends traditional library functions and resources into a purely web-based searchable system. The reason why this research has chosen digital libraries is two-fold: 1) First, most digital libraries are designed to support users to apply both searching and browsing strategies (Shen et al., 2006; Zhang et al., 2008; Huang & Xie, 2011). Digital library systems typically provide not only search functions but also browsing categories by different criteria, such as topic, subject, date, or region, in support of users' browsing activities. Users would exhibit more dynamic and various search tactics using both searching and browsing methods in using digital libraries. In search engines or online databases, users intensively rely on query searching tactics, but rarely use browsing tactics (Xie & Joo, 2009). Users of digital libraries, however, tend to apply both query searching and browsing tactics in their search processes (Kent & Bowman, 1995; Gutwin et al., 1999; Shen et al., 2006; Zhang et al., 2008; Huang & Xie, 2011). By investigating users' search behaviors in digital libraries, more various patterns of search tactics can be observed, which provide richer data in empirical analysis. 2) Second, less interactive IR research has been carried out in the context of digital libraries. In digital libraries, system designers often fail to bring appropriate system features that support

various types of search tactics compared to commercial web search engines or online databases (Xie 2008a). Digital library evaluation has been conducted in various aspects, such as collections, systems, services, administration, and others. However, there are relatively fewer studies or practices that assess user engagement and system support based on user studies. This study is one of a few studies that investigate user search process in digital library environments. It aims to contribute to the area of digital library evaluation by suggesting a research method that assesses user engagement and system support in the context of digital libraries.

1.2. Significance of the Study

This study has been motivated by one of the limitations in current IR research addressed by Järvelin (2009): *"Searchers in real-life seek to optimize the entire information access process, not just result quality. Evaluation of output alone is insufficient to explain search behavior."* (Järvelin, 2009, p. 1) IR studies have benefited greatly from system-driven experimental studies such as the Cranfield tests, Text REtrieval Conference (TREC), Cross-language information retrieval (CLEF) and NII Testbeds and Community for Information access Retrieval (NTCIR) that added our knowledge of how to create test collections, measure the quality of search results, and interpret the results of an algorithmic evaluation (Borlund & Ruthven, 2008). Interactive IR began to explicitly include users, tasks and associated context in dynamic settings of IR evaluation. Interactive IR researchers came up with more realistic evaluation criteria, such as utility and usefulness, which better reflect users' search goals and objectives (Belkin et al. 2008;

Belkin et al. 2009; Cole et al. 2009). However, interactive IR evaluation is still concerned with primarily assessing how relevant or valuable the retrieved results are to the user's search goals. Research on interactive IR has been exclusively focused on the quality of search outputs (Järvelin, 2009). Fewer studies tried to assess user search tactic patterns and the quality of search process as a method to evaluate information retrieval system performance.

As efforts to understand information search processes, many information user behavior models were developed, such as Berry-picking (Bates 1979), Information Search Process (ISP) model (Kuhlthau 1991), Information Seeking Episode Model (Belkin 1993) and Planned-situational Interactive IR model (Xie 2008a). These models have greatly contributed to the understanding of complex, dynamic nature of information seeking behavior focusing on search process. However, these models are created to conceptually understand information searching processes, rather than quantitatively predict user search patterns and assess the quality of search process. Few information seeking models include thorough quantitative examination of users' search tactic application in the search process.

In addition, there is little research in regard to evaluation criteria and measurement in interactive IR, which makes it difficult to conduct evaluation studies in interactive IR research (Kelly, 2009). Interactive IR is concerned with various users' activities, thoughts and feelings, and associated contextual factors. Especially, cognitive, affective and contextual variables are often hard to observe and to represent using categorical or numerical scale. Thus, identifying feasible measures is one of the challenging problems that interactive IR research currently encounters (Belkin et al., 2009).

In an effort to go beyond the limitations of existing IR research, this dissertation addresses the need for a study that analyzes users' search tactic application and assesses the quality of search process in using digital libraries.

First of all, interactive IR research needs to include the evaluation of the search process that consists of sequential deployment of search tactics as well as search outputs. Along with the quality of search results, it is important to assess how well search tools support users to achieve their search goals during a search process (Vakkari & Kekäläinen, 2011; Vakkari, 2013). The assumption that supporting search process, in particular search tactics, leads to the improvement of an IR system underlies this study. Belkin, Marchetti and Cool (1993) noted that the goal of IR systems is to support a range of information-seeking behaviors during the search process. This study strives to assess the search process by investigating users' search tactic application and associated system support in using digital libraries. Moreover, this study attempts to empirically examine the assumption that users' search tactic application and corresponding system support would be related to the quality of search outputs.

In addition, interactive IR evaluation needs to cover various types of users' search strategies and tactics comprehensively. Many studies have demonstrated that users do engage in a variety of different search behaviors during a search session or a search episode (Marchionini, 1995; Cool & Belkin, 2002; Olston & Chi, 2003; Lin & Xie, 2013). Researchers have pointed out that different search strategies require different kinds of system assistance to accomplish a certain search goal (Bates, 1990; Belkin et al., 1995). This study explores different types of search tactics simultaneously while investigating search sessions.

1.3. Research Questions

This dissertation intends to understand how users engage in a search session and how the system supports users' engagement in using digital libraries. In addition, this study examines the causal relationships between search processes and IR outputs. First, to explore user engagement, this study investigates how users apply search tactics while conducting different types of search tasks. Frequency, time spent, and patterns of search tactic applications are measured to understand user engagement in the search process. Second, this study identifies different types of system supports for search tactic application, and measures the degree of system support for search tactic application. Additionally, users' perceived difficulty and satisfaction with search tactic application are measured as a way to evaluate the quality of search process. Third, the author examines the causal relationships between users' search tactic application and search outputs. This study is designed to address the following research questions:

- 1) How do users engage in a search process by applying different types of search tactics while conducting different search tasks?;
- 2) How does the system support users to apply different types of search tactics?;
- 3) How do users' search tactic application and system support for different types of search tactics affect search outputs?;

RQ 3 has specific hypotheses as it statistically tests causal relationships between users' search tactic application and search outputs. Detailed hypotheses are described in Chapter 3.

1.4. Design of the Study

To answer the research questions, a user study was conducted with sixty users of digital libraries. The user study explored users' application of search tactics (RQ 1), identified different types of system support for search tactics, measured perceived system support, difficulty, and satisfaction with search tactic application (RQ 2), and examined the effects of users' search tactic selection on search outputs (RQ 3). Sixty student subjects representing different disciplines from a state university were recruited for the empirical study. They conducted three pre-defined search tasks – known-item searching, specific information searching, and exploratory searching – using Library of Congress Digital Libraries. Multiple data collection methods were used, including questionnaires, transaction logs and think-aloud protocols. Collected data were analyzed both qualitatively and quantitatively. Qualitative analysis was used to identify different types of system support for search tactics. Quantitative analysis was used to explore users' search tactic application, and measure system support for different types of search tactics. In addition, inferential statistics was used to examine the relationships between search tactic selection and search outputs. In quantitative analysis, both descriptive analysis and inferential statistics were applied, ranging from descriptive statistics, kernel regression, repeated-measures ANOVAs, hierarchical clustering, correlation analysis, multiple regression, and to structural equation modeling (SEM).

1.5. Dissertation Overview

In order to answer the established research questions, the author first examines a wide range of relevant literature to acquire theoretical bases and to set up a conceptual research model of the study (Chapter 2). Data collection and analysis methods are designed, and relevant variables are operationalized in the methodology chapter (Chapter 3). Then, the results of the analysis corresponding to each of research question are discussed in three aspects, including users' search tactic application patterns (RQ1), the assessment of system support, difficulty and satisfaction (RQ2), and the effects of search tactic application on search outputs (RQ3) (Chapter 4). Finally, the author discusses users' unique search tactic application and associated system design implications in the context of digital libraries (Chapter 5 & 6).

1.5. Chapter Summary

This chapter highlights a new study design that investigates user engagement and system support in searching digital libraries. This study is motivated by the fact that there has been relatively less research investigating the evaluation of search process in the area of IR. Also, this study tries to quantitatively model search process by analyzing users' search tactic application. Three research questions are proposed to investigate user engagement and system support in digital library systems.

CHAPTER 2. LITERATURE REVIEW

In this chapter, previous literature was reviewed in relation to user engagement and system support, search tactics, IR system evaluation, interactive IR, and information seeking process models. Also, the author addresses some limitations of existing studies, and then provides a research model of this dissertation at the conceptual level.

2.1. Definitions of Major Terms

To begin with, key terms and concepts are defined in relation to digital libraries, information seeking and behavior, and interactive information retrieval.

In this dissertation, *digital libraries* are defined as "*representations of emergent and complex forms of digital information organization and design, consisting of multiple layers and building blocks, in various stages of development* (Matusiak, 2010, p.15)."

This study is to develop an interactive IR model to assess user engagement and system support in using digital libraries. This study measures several different aspects of search process, such as system support, difficulty, and satisfaction. A *measure* is "a unit of scale to the determination of the magnitude of a quantity (Scrivin, 1991, p. 226)." In this dissertation, *IR systems* indicate any computer-mediated information resources, such as search engines, websites, online databases and digital libraries, which interact with information users to convey information responding users' requested needs (Xie 2008a).

In interactive IR, interaction between user and system or user and information is the main concern of research. In information science, interaction is a sort of subordinate

conception of information behavior in a large sense. According to Wilson (1999), *information behavior* is a broad concept that covers activities a person may engage in when identifying his or her own needs for information, searching for such information, and using or transferring the information. *Information seeking* is a subset of information behavior and can be seen as a person's purposive seeking for information in order to satisfy an underlying goal (Wilson, 2000). *Information retrieval* has an even more narrow focus and concerns formal algorithmic processes of representation, storage, searching, finding, filtering, and presentation of potential information perceived relevant to a requirement (Ingwersen & Järvelin, 2005, p. 21).

Interactive IR is a discipline concerned with users in the process of directly consulting an IR system (Robins, 2000). According to Ingwersen & Järvelin (2005), *interactive IR* refers to "the interactive communication process that occurs during retrieval of information by involving all major participants in information seeking and retrieval (p.21)." Their definition of interactive IR emphasizes communication between all related participants (e.g., user, intermediary, system) in the IR process. This reflects the transition in IR evaluation paradigm by incorporating interactions between different stakeholders and the system.

Interactive IR lies between information seeking and information retrieval. Interactive IR extended traditional system-driven IR by adopting users' cognitive aspects and contextual factors while maintaining controlled experimental design and search results evaluation in classic IR evaluation. Skov (2009) described this hierarchical structure in a nested diagram model, and the researcher modified Skov's nested diagram by incorporating representative contextual factors in interactive IR as shown in Figure 2-1.

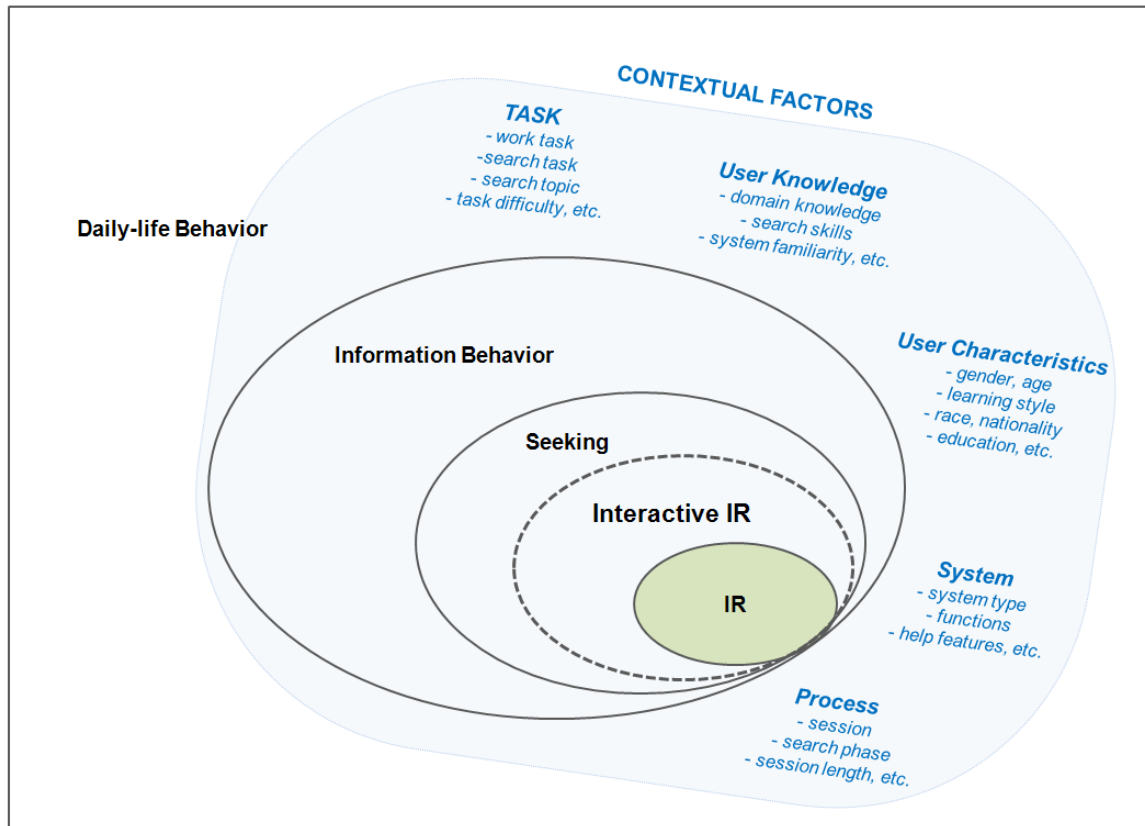


Figure 2-1. A nested model of information behavior, seeking, interactive information retrieval, and information retrieval – Extension of Skov's diagram (Skov, 2009, p.18).

Kelly (2009) differentiates interactive IR from both system-focused traditional IR and human-focused information seeking behavior studies (Figure 2-2). While system-driven IR evaluation takes humans out of the evaluation model, interactive IR incorporates users' behavior and experiences and the interactions that occur between users and systems and users and information (Kelly, 2009). In interactive IR evaluation, represented by TREC Interactive Track, a system or interface feature is typically being evaluated, and humans are directly involved in evaluation in terms of human behavior and cognition and information seeking context. This study lies around "Information Seeking Behavior with IR systems" as the nature of this study is exploratory. At the same time, this study involves some output measures that are discussed in TREC interactive style studies.

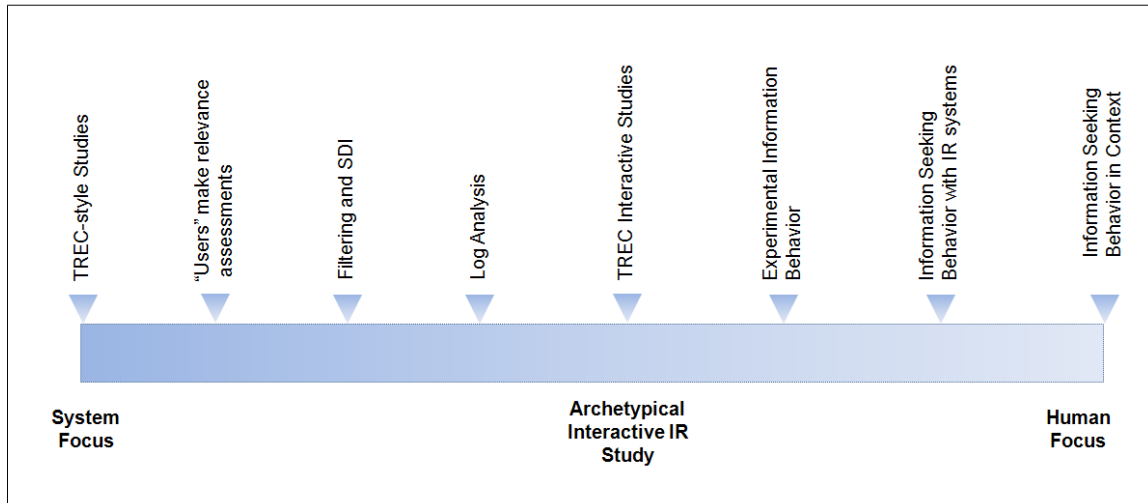


Figure 2-2. Research continuum for conceptualizing interactive IR research (Kelly, 2009, p.10)

In this study, two important concepts are *system support* and *user engagement*. The definition of *system support* is the representation of system features to assist users' behavioral activities or cognitive intention and to facilitate user-system interactions during an IR process. *User engagement* is users' behavioral and cognitive intention to change the IR process and responses to the outcome from the IR system (Xie, 2003).

User engagement is represented by users' selection of *search strategies* and *search tactics* in the IR process. When users engage in the IR process, they must have some search strategies that are a combination of the choice of search tactics (Vakkari et al., 2003). A *search tactic* is a move or moves made to further a search, and a *search strategy* is combinations of search tactics applied to accomplish information search tasks as a plan for the whole search (Bates, 1979; Xie 2008a). Table 2-1 summarizes key concepts and their definitions in this study.

Table 2-1. Definitions of key terms and concepts

Concept	Definition	References
Digital libraries	Representations of emergent and complex forms of digital information organization and design, consisting of multiple layers and building blocks, in various stages of development.	Matusiak (2010)
Measure	A unit of scale to determination of the magnitude of a quantity.	Scrivin (1991)
Information behavior	A broad concept that covers activities a person may engage in when identifying his or her own needs for information, searching for such information, and using or transferring the information.	Wison (1999)
Information seeking	A subset of information behavior and can be seen as a person's purposive seeking for information in order to satisfy an underlying goal	Wilson (2000)
Information retrieval	Formal algorithmic processes of representation, storage, searching, finding, filtering, and presentation of potential information perceived relevant to a requirement	Ingwersen and Järvelin (2005)
Interactive information retrieval	The interactive communication processes that occur during retrieval of information by involving all major participants in information seeking and retrieval.	Inwersen and Järvelin (2005)
System support	Representation of system features to assist users' behavioral activities or cognitive intention and to facilitate interactions in IR processes.	Xie (2003)
User engagement	A user's behavioral and cognitive intention to change the IR process and responses to the outcome of the IR system. <i>User engagement</i> is represented by users' selection of <i>search strategies</i> and <i>search tactics</i> in the IR process.	Xie (2008a)
Search strategy	The products of planned or situational interactions between users and IR systems	Bates (1979), Xie (2008a)
Search tactic	User action related to search process to achieve specific objectives in the information searching process	Bates (1979), Xie (2008a)

2.2. User Engagement and System Support

In this section, the author surveys previous literature that explored the relationships between user engagement and system support, which are the main concepts to be investigated in this dissertation.

The nature of IR is interactions between user engagement and system support during the search process. A search session can be accomplished by the interactive process where users physically and intellectually engage in different system features and information. Interactive IR process can be viewed as how a user balances two different dimensions of labor, namely user engagement and the system's reactive support in terms of: 1) to what extent a user should exert his/her effort to control the IR process and 2) to what extent the system should support the user to easily proceed to the IR process (Bates, 1990; Xie & Cool, 2000; Xie, 2003). In this study, these two key concepts are defined as:

- *User engagement* refers to a user's behavioral and cognitive intention to change the IR process and responses to the output of the IR system (Xie, 2003; Xie 2008a).
- *System support* refers to representations of system features to assist users' behavioral activities or cognitive intention and to facilitate their interactions with the system in IR process (Xie, 2003).

Bates (1990) is one of the early scholars who started a provocative discussion of balancing a system role and a user role in regard to IR system design, "what capabilities should we design for the system, and what capabilities should we enable the searcher to exercise?" This suggests that IR system design is a matter of dividing labor between user

and system, and we should consider both sides of user and system roles. When developing an IR system, we need to think about to what extent users should engage in the search process and to what extent the system helps their engagement. Users' engagement can be optimally supported by different combinations of IR features (Bates 1990). Belkin (1993) paid attention to user roles in IR. He emphasized the importance of user participation in the process of IR. In the IR process, users are requested to be an active participant rather than a passive recipient of and reactor to output from the IR system (Belkin, 1993).

Several researchers discussed the importance of user roles in designing IR systems. Hendry and Harper (1997) addressed the problems of over-determined systems and suggested IR system interfaces that users have more control in solving information problems. Savage-Knepshield and Belkin (1999) reviewed IR system design trends, and claimed that IR system designers increased the level of control provided to users as well as system support features. As users have more engagement, IR systems afforded more dynamic interactions between users and systems. Xie and Cool (2000) conducted an evaluation study of online databases. Their evaluation study is based on Bates's premise comparing roles of system support and user engagement in IR process. They assessed functionalities of online databases, and yielded several IR system design implications that support both ease of use and user control to satisfy diverse needs of both novice and experienced users.

Xie (2003) first attempted to directly compare users' perceptions of ease-of-use versus user control. The major finding of that study is that the level of system support differs by system feature. For example, when using search limiter function, users experienced more

support for user control than ease-of-use. She also explored desired functionalities and interface structures of IR systems in supporting of both ease-of-use and user control. She concluded ease-of-use and user control are two essential factors necessary to lead to effective retrieval. Also, she proposed a model of optimal support for ease of use and user control that describes what are the system roles and user roles in conducting various IR sub-tasks (Figure 2-3).

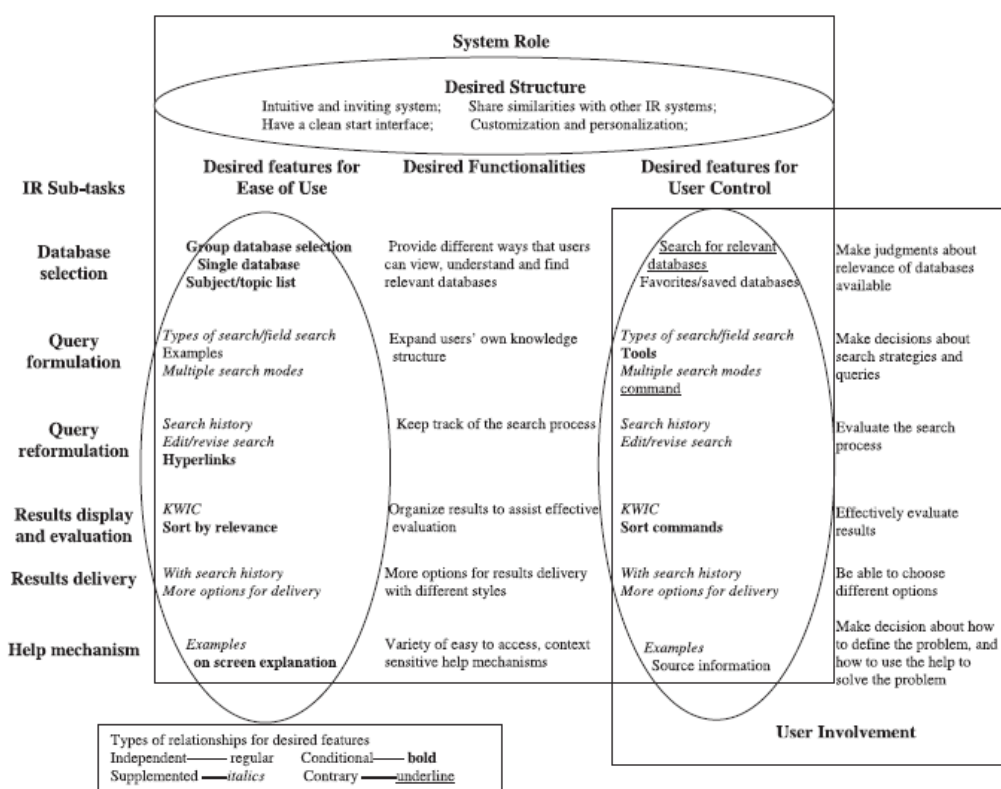


Figure 2-3. Model of optimal support of ease of use and user control (Xie, 2003, p. 916)

Recently, Marki et al. (2008a) matched types of search tactics and related system support features. For example, they matched segmented search fields with query searching, while document metadata were suggested for users' resource selecting tactics. In addition,

Marki et al. (2008b) proposed an information system evaluation method that covers both user and system aspects. They tried to assess both usability and functionality based on Ellis's information seeking model. In their method, usability evaluation examines how easy it is to use specific system features while functionality evaluation aimed at system support for users.

2.3. Search Tactics

In this study, the unit of analysis is the search tactic. The definition of a search tactic is a move or moves made to further a search process by achieving a specific subtask during the search process (Xie 2008a). Search tactics have been recognized as a means of examining the search process at a micro level (Yue, Han & He, 2012). A search tactic is the most granular level of user action that is needed for users to accomplish a sub-task in a search process. A search tactic consists of a single move or moves that involve user intention to complete a specific objective required to proceed with a search process.

Bates (1979) is one of the early scholars who identified different levels of user behaviors in information seeking processes. She defined four hierarchical levels of taxonomy in relation to information seeking:

- Search tactic: A move made to further a search.
- Search strategy (in searching): A plan for the whole search.
- Search strategy (as an area of study): The study of the theory, principles, and practice of making and using search strategies and tactics.

- Search behavior: What people do and/or, as far as can be determined, what they think when they search.

Bates classified search tactics into four dimensions: 1) monitoring tactics refers to keeping aware of the overall progress of the search (e.g., check, correct, record, etc.); 2) file structured tactics refers to using the structure of the database to find information (e.g., browse, select, survey, etc.); 3) search information tactics indicate an aid to the formulation of specific search commands (e.g., specify, exhaust, reduce, etc.); and 4) term tactics are to aid in the selection and revision of specific terms within the search formulation (e.g., super, sub, relate, etc.).

Smith (2012) further extended Bates' structure of search tactics into the Internet environment. He identified 34 search tactics in five dimensions by extending Bates' original 29 tactics. In his search tactic identification, he added the dimension of evaluation tactics, and newly identified specific tactics, such as context evaluation, crosschecking, and audition. His research contributed to the reinterpretation of Bates' information search tactics reflecting the unique context of internet searching, and further suggested search tactics related to evaluation behavior.

Marchionini (1995) identified different sequential sub-tasks in information seeking process. He defines information seeking as a process in which information seekers purposefully engage to change their state of knowledge. His model proposes eight sub-tasks during an information seeking process (Figure 2-4): (1) recognizing and accepting an information problem; (2) defining and understanding the problem; (3) choosing a search system; (4) formulating a query; (5) executing search; (6) examining results; (7)

extracting information; and (8) reflecting/iterating/stopping. According to his model, information seeking begins with the recognition and acceptance of a problem and continues until the problem is resolved or abandoned. The information problem can be internally or externally motivated, and it can be characterized as a gap, a visceral need, or an anomaly. In the initial sub-task, a user becomes “aware” of the problem. Once the user defines the search problem, he/she is required to choose an adequate search system. In this sub-task, user knowledge and experiences strongly influence the selection of a system. Query formulation is a sub-task that matches understanding of the task with the system selected. The initial query defines an entry point to the system and is followed by browsing and/or query reformulation. Search execution is related to physical actions to the system. Users employ different execution techniques for different systems. Search results examination refers to the user’s assessment of the response from the search system, and the relevance assessment leads to information extraction. Using different skills, such as reading, scanning, copying, and storing, users manipulate and integrate obtained information into their knowledge of the domain. An information search is usually completed with the iteration of sub-tasks, and a user monitors his/her search process and determines when to stop dependent on both internal and external functions. In his model, Marchionini not only comprehensively defined multiple sub-tasks during the search process but also explored transitions in those sub-tasks.

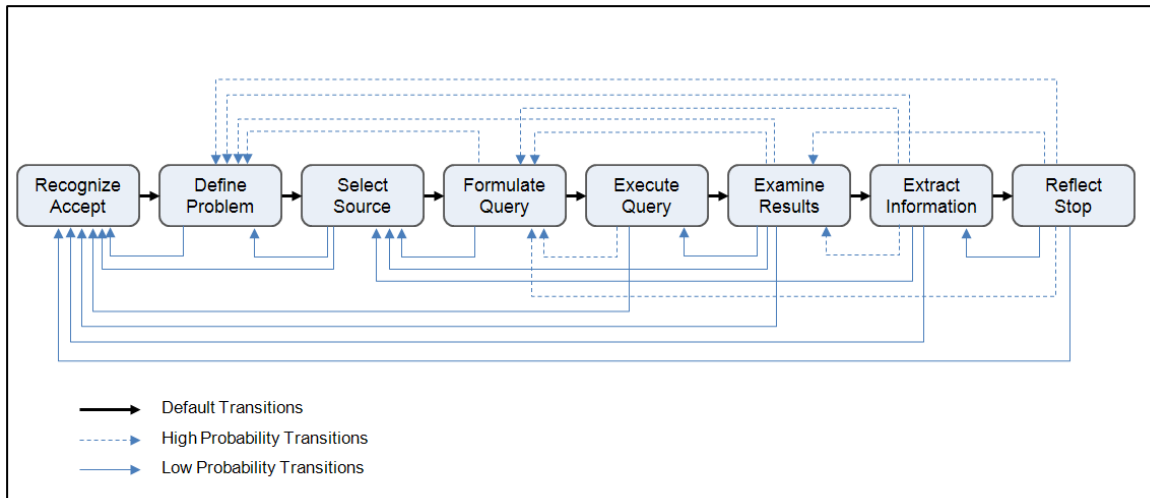


Figure 2-4. Marchionini's Information Seeking-Process Model (Marchionini, 1995, p.50)

Xie and Cool (2000) identified six sub-tasks in which users would engage during the IR process to accomplish their search tasks in searching online databases: (1) databases selection, (2) query formulation, (3) query reformulation, (4) help mechanism access, (5) results organization and display, and (6) results delivery. In her planned-situational interactive IR model, Xie (2003; 2008a) identified twelve interactive intentions that lead the search process to accomplish search goals. An interactive intention refers to a micro-level sub-goal that a user has to achieve to accomplish his/her current search goals, and those intentions are the products of plans and situations. Those twelve interactive intentions are: identify, learn, explore, create, modify, monitor, keep records, access, organize, evaluate, obtain, and disseminate. Xie and Joo (2010a) further extended Xie's (2008a) identification of interactive intentions. They tried to come up with a model to describe IR processes at the micro-level based on the analysis of search tactic transitions.

Recently, Kules and Capra (2011) explored different stages of user engagement in exploratory search sessions. They identified five different stages during an exploratory

search session, including Query, Overview, Extracting, Deciding Next, and Deciding Topic. Yue, Han and He (2012) explored the relationship between users' actions and search tactics based on Hidden Markov Model (HMM). They assumed that user search actions include latent user search tactics, and examined the transitions of search tactics using HMM algorithm. They identified five distinct search tactics, including Query, View, Save, Workspace, Topic and Chat. Then, they related these five user actions with sub-process, such as defining problem, selecting sources, and examining results.

2.4. Information Retrieval Evaluation

2.4.1. System-oriented IR System Evaluation

In IR evaluation, system-oriented studies have focused on the development and evaluation of effective representation techniques, storage, matching algorithms and indexing techniques to be implemented in IR systems. The first evaluation effort of IR systems dates back to the late 1950s and early 1960s. At that time, Cleverdon and his colleagues at Cranfield College of Aeronautics investigated indexing languages for IR (Cleverdon, Mills & Keen, 1966; Cleverdon, 1967). Cranfield Tests have set the classic paradigm of research on IR system evaluation. The original purposes of Cranfield tests were to examine indexing systems for IR. In the tests, four forms of indexing systems, including universal decimal classification (UDC), alphabetical subject heading, faceted classification, and Uniterm system, were compared using the test collection of 18,000 documents and 1,200 search terms (Cleverdon, 1991).

Cranfield tests built up the principle of test collections including a document corpus, a set of information requests, and relevance judgements. This setup supports experimental control of variables and comparability of results and influenced the inception of the Text REtrieval Conference (TREC). TREC IR evaluation has been upgraded by applying new types of data collections, enhancing evaluation settings, and scaling up to very large test collections. Figure 2-3 delineates the system-driven IR evaluation, which consists of a database, algorithms, requests and relevance assessments. The main focus of the system-oriented IR research is on matching between document representation and request representation (Ingwersen & Järvelin, 2005).

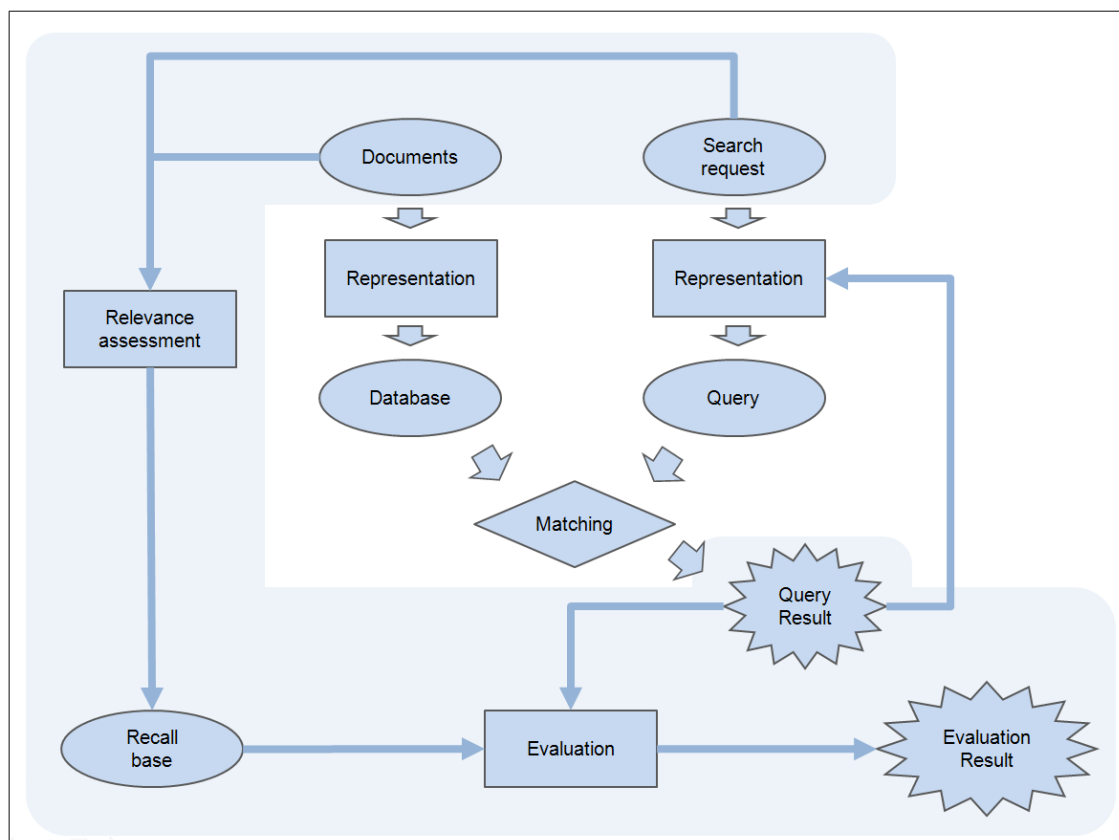


Figure 2-5. The basic laboratory model of IR (Ingwersen & Järvelin, 2005, p.115)

The two fundamental effectiveness measures are *recall* and *precision*, which have been most widely applied in IR performance evaluation to date. *Precision* indicates the proportion of retrieved documents which are relevant, and it is a measure of accuracy of the results. *Recall* refers to the proportion of relevant documents in the collection which have been retrieved (Harman & Voorhees, 2006). *F-measure*, which is a harmonic mean of precision and recall, is a frequently used summary measure for IR performance

$$(F = \frac{2 * precision * recall}{precision + recall}).$$

In addition, several variants of precision and recall measures have appeared in TREC conferences. For example, *precision and recall at different cut-off level* uses ranking information of retrieved documents (Voorhees, 2003). A cut-off level of 100 documents indicates the subset of top 100 ranked document in the total retrieved set. This variant is used to evaluate the results in the situation that only top retrieved documents are considered in the tests. It underlies the assumption that most users usually do not look at the results beyond the third page in result lists. *Mean Average Precision* is another

summary measure of a ranked retrieval run ($AveP = \frac{\sum_{r=1}^N (P(r) \times rel(r))}{\text{Number of relevant documents}}$). It is

calculated as the average of precisions at the point of each of the relevant documents in the ranked sequence. This measure is designed to weight documents retrieved in higher rankings more heavily than documents retrieved in lower rankings (Turpin & Scholer, 2006). Some IR performance measures are focused on search efficiency. For instance, Kaki (2004) introduced two unique evaluation measures to represent the efficiency of IR systems, namely *Search Speed* and *Immediate Accuracy*. *Search Speed* is an attempt to measure answers per minute, and it is computed easily by dividing the number of answers

found by the time it took to find them ($Search\ Speed = \frac{Answers\ found}{Minutes\ searches}$). *Immediate*

Accuracy is to capture the success of web search. It is calculated as the proportion of cases where a user obtains at least one relevant result in the n^{th} result selection

$$(immediate\ accuracy_n = \frac{number\ of\ first\ relevant\ results_n}{total\ number\ of\ observations}).$$

System-centered methods have significantly contributed to IR system evaluation, especially to retrieval algorithm, test collections, indexing, and quantifiable IR performance measurement. However, relevance judgment, which is indispensable in precision and recall, still remains as a controversial component in IR research due to its complex nature of relevance (Saracevic, 2007). As collections in TREC became huge and relevance judgment of documents is needed to be sophisticated, consistency in relevance judgment among human assessors could be a critical factor on the reliability and validity of TREC (Aslam *et al.*, 2006). In addition, system-oriented approaches did not include users and context in the experimental design of IR evaluation. The traditional IR laboratory setup, which mostly consists of comparison or representation, eliminates the human searcher from the experiment to control for variables and ensures that the effects found in the research are due to variations in system parameters (Beaulieu, Robertson, & Rasmussen, 1996). System-oriented approaches make IR research challenging from the cognitive side, particularly in interactive environments (Tamine-Lechani *et al.*, 2009). Moreover, traditional IR does treat information needs as a static concept entirely, and uses mostly binary, topical relevance while ignoring the fact that relevance is a multidimensional and dynamic concept (Borlund, 2000). In particular, precision and recall as an indicator for successful retrieval provide only summarized average results to the detriment of a deeper understanding of search processes (Ingwersen & Järvelin, 2005).

Even though precision and recall have been used as the most popular measures, users may not always care about precision or recall rate in IR system performance (Su, 1992; Hearst, 2006). For example, users might be satisfied with if they find appropriate information for their task, instead of finding unnecessarily many relevant items.

2.4.2. User-oriented IR System Evaluation

The opposite pole of system-centered IR evaluation in information seeking and retrieval research is user-centered evaluation. User-side IR evaluation studies have been conducted parallel to system-oriented IR in an attempt to assess the system in the perspective of user experiences. User satisfaction is one of the most widely applied evaluation criteria in the evaluation of IR systems directly from users' claims (Hert, 2001). User satisfaction has been considered an obvious way to judge the fit of an IR system with user tasks, and it checks how users are satisfied with the system by directly asking to those who used the system. Satisfaction is generally acknowledged as a prime criterion for judging the quality of information system performance (Brophy, 2006). Su (1992) pointed out three potential advantages of the user satisfaction criterion in IR evaluation. User satisfaction takes explicit account of users and their subjective evaluation of various aspects of IR interactions; focuses on multi-dimensional evaluation of the IR process; and recognizes users and their request characteristics as possible factors that influence user evaluation (Su, 1992).

In the field of information science, however, user satisfaction has been controversial in IR evaluation. Belkin and Vickery (1985) addressed ambiguity in the definition of satisfaction and a possible validity problem caused from its subjective nature. Also,

Hildreth (2001) questioned about the reliability of a satisfaction measure due to its lack of independence from other influential factors during the procedure of information seeking.

When it is used as a measure for IR system evaluation, it can be easily affected by non-performance factors that may confound the results (Al-Maskari & Sanderson, 2010).

Moreover, findings from satisfaction measurement usually do not offer direct system design implications. Despite of these potential issues, many researchers adopted user satisfaction as one of the key measures to represent the quality of information systems.

In user side evaluation of IR systems, usability evaluation has been widely conducted in both research and industry fields. Usability evaluation is a broader approach that covers multiple aspects of IR systems, such as effectiveness, efficiency, learnability and satisfaction, from the perspective of user experience with an IR interface. Usability evaluation attempts to answer questions like: “Can people use this system?” or “Does the system help people do their jobs better?” (Morse, 2002).

In general, usability consists of multiple attributes such as efficiency, effectiveness, satisfaction, learnability, errors, among others. Nielson (1993) is one of the representative experts in the field of usability engineering. According to Nielson, usability is defined as “*a narrow concern compared to the larger issue of system acceptability, which basically is the question of whether the system is good enough to satisfy all the needs and requirements of the users and other potential stakeholders, such as the users’ clients and managers* (Nielson, 1993, p. 24).” In his model, usability is comprised of five subordinate attributes (Nielson 1993, p.25):

- *Learnability*: The system should be easy to learn so that a casual user can begin to work quickly using it.
- *Efficiency*: The system should be efficient to use so that a casual user can yield high productivity when he or she is already accustomed to use it.
- *Memorability*: The system should be easy to remember so that the casual user is able to return to the system after some period of not having used it, without having to learn everything all over again.
- *Errors*: The system should have a low error rate, so that users make few errors during their use of the system and can easily recover from any error they may make.
- *Satisfaction*: The system should be pleasant to use so that a user can feel the subjective satisfaction from using it.

Another widely cited usability evaluation model is ISO 9241-11, which is an international standard by International Standards Organization (ISO, 1997). According to ISO 9241-11, usability refers to “the extent to which a product can be used by specified users to achieve specified goals with (1) effectiveness, (2) efficiency and (3) satisfaction in a specified context of use (ISO 1997, p. 2).”

- *Effectiveness*: Accuracy and completeness with which users achieve specified goals.
- *Efficiency*: Resources expended in relation to the accuracy and completeness with which users achieve goals.

- *Satisfaction*: Freedom from discomfort, and positive attitudes towards the use of the product.

Additionally, many researchers have proposed a variety of usability attributes in their own usability studies (Table 2-2). However, usability tests purely concern the interface design, not the IR performance or process. The main purpose of a usability test is to diagnose errors and to find functional problems in the interface and to conclude interface design implications to enhance the effectiveness and efficiency of the system. In some aspects, a usability test shares some common features with interactive IR evaluation in that usability tests are also based on task-driven evaluation. In user-participatory usability evaluation, participants are asked to conduct a series of predefined system use tasks. The difference lies in that a usability test focuses on functionality and outcomes, while interactive IR investigates more dynamic system performance, user-system interactions, as well as IR outcomes. A usability test points out the weakness or problems of a system in overall interface design, but it does not tell in which search processes users would get appropriate support or feel difficulty in detail.

Table 2-2. The comparison of usability attributes by different researchers

Researchers	Attributes
Nielson (1993)	Learnability, Efficiency, Memorability, Errors, Satisfaction
ISO (1998)	Efficiency, Effectiveness, Satisfaction
Jeng (2005; 2006)	Efficiency, Effectiveness, Learnability, Satisfaction
Shackel (1991)	Learnability, Flexibility, Effectiveness, User attitude
Brink, Gergle, Wood (2002)	Easy to learn, Efficient to use, Easy to remember, Error tolerant, Functionally correct, Subjectivity, Pleasing
Hix & Hartson (1993)	Initial performance, Long-term performance, Learnability, Retainability, Advanced feature usage, First impression, Long-term user satisfaction
Oulanov & Pajarillo (2002)	Affect, Efficiency, Control, Helpfulness, Adaptability
Ward, Hiller (2003)	Time and effort, Context, Accomplishment, Users' reaction
Joo & Lee (2010)	Efficiency, Effectiveness, Learnability, Satisfaction

2.5. Interactive IR Evaluation

Interactive IR evaluation can be viewed as an alternative approach to overcome the limitations of system-driven IR, satisfaction evaluation, or usability evaluation.

Interactive IR evaluation framework offers a comprehensive view concerning system performance, usability, and interactive process simultaneously (Xie, 2008a). The incorporation of user interactions into IR system evaluation has been important concerns for IR researchers (Callan et al. 2007).

2.5.1. TREC Interactive Track

The TREC Interactive Track, which lasted for nine years, made important contributions to IR performance evaluation incorporating user engagement and tasks. In TREC Interactive Track, IR system design directly involves users in evaluation experiments using predefined search scenarios and tasks. The focuses of this track were to investigate 1) the searcher's interaction with the IR system; 2) the searcher's behavior, search process, and interim results as well as final results; 3) the effects of system, topic, searcher, and their interactions; and 4) the assessment of the evaluation methodology (Over 2001). For nine years of TREC Interactive Track, experimental search tasks and topics evolved to reflect more dynamic and diverse user-system interactions during the IR process.

The initial Interactive Track was formed from TREC-3. In the first Interactive Track, several participant groups recruited human subjects to compare manual routing query construction to the fully automated routing system. Subjects were asked to create optimal routing queries using training data to fifty standard routing topics and participants experimented with a variety of system features to support subjects in achieving the tasks. Major findings reveal that human created routing queries were less favorable compared to automated ones, in that automated systems used more training data and were better at capturing collection distribution data (Over, 2001; Hersh, 2009). In TREC-4 Interactive Track, twenty-five ad-hoc search tasks were adopted instead of routing tasks, and subjects were asked to retrieve as many documents as possible in 30 minutes and construct the best query. This Interactive Track compared the results of interactive ad hoc searches with automatic searching and found a difference of relevance assessment

between subjects and assessors (Dumais & Belkin, 2005). In TREC-5, the Interactive Track introduced a new type of search task, so called *aspectual recall task*, which required subjects to find as many relevant documents as possible that discussed different aspects of a topic. The TREC-6 Interactive Track used same searcher tasks to TREC-5 with six ad hoc topics, and used aspectual precision, which refers to the fraction of documents saved containing at least one aspect, along with aspectual recall. The main analysis confirmed statistically significant main effects of topic, system and searcher. In TREC-7 Interactive Track, cross-site comparison was dropped because it was difficult to have a direct comparison considering the requirements of the track. Also, the term “aspect” was replaced by “instance” in order to make searchers easier to understand (Hersh & Over, 2000). TREC 9 focused more on fact-finding tasks in an effort to reduce the length of experimental sessions and examine more tasks and collection. In TREC10 and11, Web searching was conducted by participants and Web-track collection was used as a common collection (Hersh & Over, 2003).

The TREC Interactive Track contributed to the development of a common framework for evaluation and comparison of interactive IR systems by offering applied methodologies, experimental designs, and techniques for results reporting (Dumais & Belkin, 2005). However, the Interactive Track was based on fixed search task, topics and collection, and judgments from selected assessors. Thus, strictly speaking, it failed to represent real interactive IR processes. The evaluation method used in the Interactive Track was not flexible enough to investigate IR processes in real environments (Ingwersen & Järvelin, 2005).

2.5.2. Evaluation Criteria in Interactive IR

There is a need for a user-centered perspective on relevance as relevance judgment is subjective in nature and involves individual mental experience (Borlund, 2003). It is needed to come up with more diverse relevance judgment evaluation to better reflect users' complicated needs and tasks. In interactive IR, relevance judgment does not rely on binary measurement of relevance any more. Precision and recall are insufficient for evaluating interactive IR systems because those cannot quantify how much search results would be informative to resolve a search task (Cheng, Hu, & Heidorn, 2010). Different levels of relevance became an important concept in interactive IR, such as partially relevant and situational relevance (Spink & Greisdorf, 2000). For example, Schamber, Eisenberg and Nilan (1990) proposed a measure of situational relevance, which refers to the usefulness of an information object in relation to the work task intrinsic to the user. Borlund (2003) created alternative IR performance measures against traditional ones, namely, *relative relevance* and *ranked half-life indicator*. As these alternative measures engage in real IR situations, interactive IR evaluation deals with different levels or regions of relevance to better represent users' complicated need for the IR system. In addition, interactive IR takes into consideration the values of search results beyond simple relevance judgment. Su (1992, 1994) found that precision is not always significantly correlated with the user's judgment of success, and emphasized the importance of assessing the value of retrieved results. The value of search results usually shows a high correlation with users' satisfaction with search results, and is dependent on the tasks that users intend to accomplish.

Saracevic (1996a) further identified different aspects of relevance in IR evaluation:

- *System or algorithmic relevance*: it is to match a query and information objects (texts) in the file of a system as retrieved.
- *Topical or subject relevance*: it is to examine the relationship between the subject or topic expressed in a query, and topic or subject covered by retrieved text.
- *Cognitive relevance or pertinence*: it focuses on relation between the state of knowledge and cognitive information need of a user, and texts retrieved.
- *Situational relevance or utility*: it refers to the relationship between the situation, task, or problem at hand, and texts retrieved by the system. In particular, it focuses on the usefulness of the retrieved documents to the user's task.
- *Motivational or affective relevance*: it refers to the relationship between the intents, goals, and motivations of a user, and texts retrieved by a system.

Some researchers cared about evaluation of system support in interactive IR environments. Belkin, Cole and Liu (2009) investigated usefulness of system features in support of user-system interactions, while incorporating users' different search goals, tasks, and search strategies. They suggested a new IR evaluation model in terms of 1) how well the system supports the accomplishment of the overall task/goal; 2) how well the system supports the contribution of each interaction towards the achievement of the overall goal; and 3) how well the system supports each interaction. They offered a comprehensive viewpoint to investigate system support that facilitates user-system interactions.

In this way, interactive IR has widened search result evaluation criteria in IR research. Interactive IR researchers have considered various contextual situations and provided different IR evaluation criteria, such as aspectual recall and relative relevance. In particular, Saracevic (1996a) surveyed different aspects of relevance by examining relationships among different elements of IR processes and tasks. Additionally, Belkin and his colleagues (2009) suggested the evaluation of system features as a new approach to assess user-system interactions in IR processes.

2.5.3. Search Tasks in Interactive IR

In interactive IR research, search tasks have been considered as a key factor that influence users' search behavior. Many researchers identified various types of tasks in different IR situations. For example, researchers identified two major dimensions of tasks, including search tasks and work tasks (Ingwersen, 1992; Byström and Järvelin, 1995; Li and Belkin, 2008; Xie, 2009). Li and Belkin (2008) classified facets of tasks comprehensively as different aspects, properties or characteristics of a task. Her scheme of task facets include source, task doer, time, and others. Byström and Järvelin (1995) studied complexity of task, and five levels of complexity were identified that ranged from automatic information processing to genuine decision.

Along with the identification of task types, many researchers investigated the effects of task types on users' search behavior. For example, Hsieh-Yee (1998) examined the relationship between search tactics and search tasks. In her study, she investigated how the patterns of search tactics would differ by two types of search tasks, including known-

item searches and subject searches. She found there were no significant differences of search tactics used for these two types of search tasks. Kim and Allen (2002) investigated how search task types would influence users' search behavior. Based on experimental design, they found that users' site views and search tool uses were different between known-item searches and subject searches. Byström (2002) further examined the relationships between task complexity and information-seeking behavior. She found that task complexity has a significant impact on source uses; for example, internal official documents tend to be used more in automatic information processing task, while human sources are more selected in decision making task. Similarly, Shiri and Revie (2003) assessed the relationships between task topic complexity and search moves in using thesaurus-based IR systems. They found that more search moves are needed in more complex topic tasks. Hung (2005) also studied how search tactic selection can be influenced by task types, such as specific searches, general searches, and subjective searches. In her results, users are likely to apply complex patterns of search tactics in general and subjective searches. In her experimental study, Li (2010) found that work task affect users' search performance measured by search efficiency and effectiveness. High efficiency and effectiveness would be achieved in less complex task, and vice versa. Recently, Liu et al. (2010) investigated how task type and associated situation would affect users' query reformulation behavior. Their major findings are that three types of tasks – simple, hierarchical, and parallel searches – are related to query reformulation behavior. For example, while specialization strategies were frequently used in simple and hierarchical searches, generalization strategies were more used in hierarchical searches.

2.5.4. Digital Library Evaluation

Digital libraries became a fundamental information source for researchers, teachers, and students in support of research and education. Although digital libraries share common characteristics with other types of IR systems, it has unique features as well. Digital libraries are (1) highly dynamic and ephemeral in technical, collection and information needs (Fox & Urs, 2002); (2) highly heterogeneous along format, coverage, user, and system dimensions; and (3) tightly virtual collaboration among different groups of stakeholders, including knowledge creators, publishers, distributors, information specialists, librarians, and users (O'Day & Nardi, 2003). In the last decade, researchers and practitioners have substantially expanded their efforts to digital library evaluation. This section reviews major evaluation frameworks and associated evaluation criteria in the area of digital libraries.

The early digital library projects, funded by the National Science Foundation (NSF) as part of Digital Libraries Initiatives I and II, laid a groundwork in evaluation research by producing digital library prototypes and frameworks (Borgman et al., 2000; Buttenfield 1999; Hill et al., 2000; Van House, et al., 1996). Although these early projects weighed much on the development of digital library prototypes and models, several of them undertook evaluation studies as part of the design cycle. The Alexandria Digital Library (ADL) is one of the six digital library projects funded by NSF, DARPA, and NASA. Hill et al. (1997) used multiple methods, such as surveys, ethnographic studies, and focus groups, to obtain feedback about ADL at the University of California, Santa Barbara. Using exploratory factor analysis, they derived six dimensions for evaluating ADL: (1)

overall ease of use; (2) overall appeal; (3) terminology clarity; (4) overall usefulness; (5) overall performance; and (6) navigational clarity.

Saracevic (2000, 2004) suggested a comprehensive evaluation model that covers multiple aspects of digital libraries. Saracevic's evaluation framework consisting of six constructs is one of most widely cited models in the area of digital library evaluation research. His evaluation framework is designed to comprehensively assess multiple dimensions of digital libraries using various evaluation criteria. In particular, his digital library evaluation framework intends to assess a social contextual aspect, such as institutional fit, usefulness, sustainability, impact on community, and others. Before his evaluation framework, there was little effort that included social context in digital library evaluation. Table 2-3 presents Saracevic's digital library evaluation framework that includes six criteria and associated measures. Recently, Zhang (2010a; 2010b) further extended Saracevic's (2000) evaluation framework by adding feasible measures. She adopted evaluation criteria from Saracevic's framework, and provided very specific and feasible measures based on an empirical study.

Table 2-3. Saracevic's digital library evaluation framework

Construct	Objective
Content	how well are digital collections selected, developed; objects created, organized, represented, presented
Technology	how well do hardware & software support library functions
Interface	what is available for users to interact & how much is interaction supported or hindered
Process/ Service	what processes & assistance are provided; what range of services is available; how well are they functioning;
User	what are the outcomes of digital library use – changes in human information behavior, cognitive state, decision-making, problem-solving; impact on accomplishing tasks; broader impact/benefit in research, professional work
Context	how well does a digital library fit into, respond to, follow larger context – institutional, economic, legal, social, cultural; effects on context

DELOS model is another widely known DL evaluation framework internationally.

DELOS is a comprehensive and large scale DL project, which represents joint activities aimed at coordinating the ongoing research efforts of the major European teams working in the digital library area. Candela et al. (2007) established DELOS Manifesto that presents a three-tier digital library framework incorporating six core components such as content, functionality, quality, policy, architecture, and user. DELOS Network of Excellence has conducted a series of research concerning the evaluation of digital libraries. Fuhr et al. (2001) proposed a scheme for digital library evaluation which contains four dimensions: data/collection, system/technology, users, and usage. Tsakonas et al. (2004) further examined the interactions between digital library components and proposed the following evaluation foci: usability, usefulness, and system performance respectively. Fuhr et al. (2007) developed a digital library evaluation framework based on a DELOS model and a large-scale survey of digital library evaluation activities.

Xie's (2006; 2008b) evaluation framework has shifted a focus closer to the users, and proposed five types of criteria: usability, collection quality, service quality, system performance efficiency, and user feedback solicitation. Xie (2006) pointed out that little has been done on the identification of evaluation criteria from the perspective of users. Although many researchers proposed evaluation criteria for digital libraries, there was a lack of user input regarding evaluation criteria. Xie (2008b) surveyed users' perceptions and opinions in relation to evaluation criteria through multiple channels such as diaries, questionnaires, and interviews. Based on the results of the user surveys, she yielded a DL evaluation framework that includes a set of specific evaluation criteria. Figure 2-6 summarizes her user-driven digital library evaluation framework.

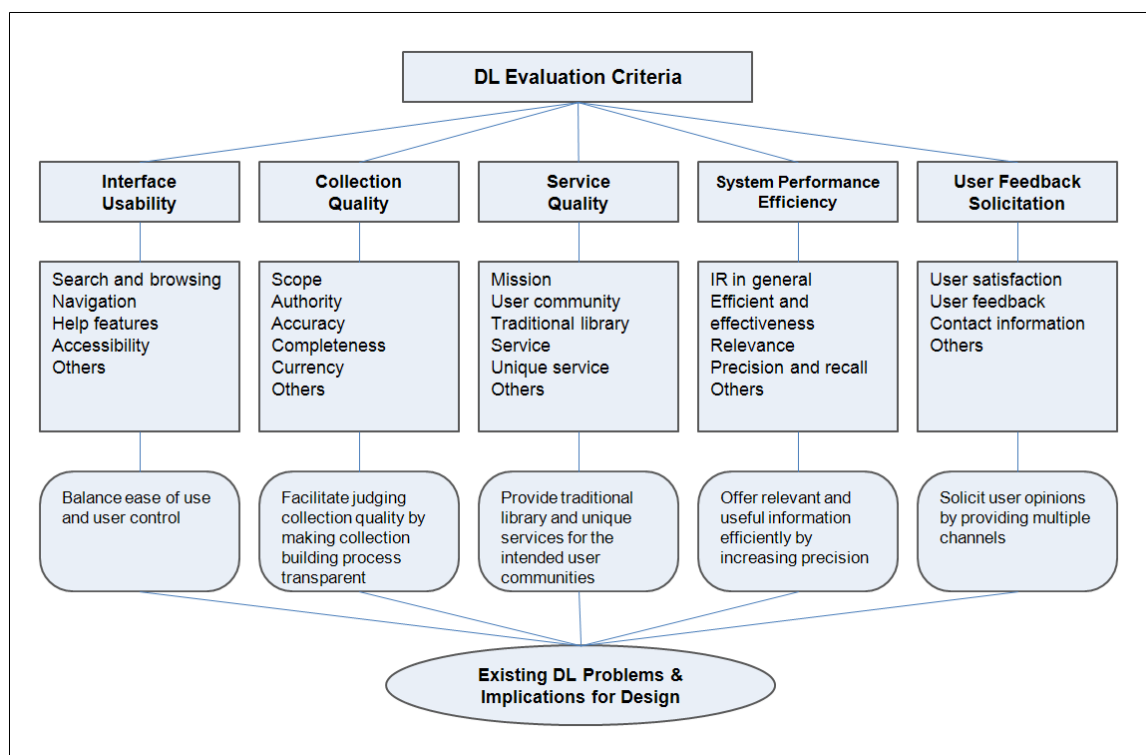


Figure 2-6. DL evaluation framework: Criteria, problem, and implications for design (Xie, 2006, p.449)

Usability evaluation is another major effort in digital library evaluation. Saracevic (2004) suggested a list of usability criteria and attributes customized to DL evaluation, such as content, process, format and overall assessment. Dillon (1999) proposed a qualitative framework (TIME) for DL usability evaluation that covers user task (T), information model (I), manipulation facilities (M), and the ergonomic variables (E). Ward and Hiller (2005) suggested usability evaluation criteria specific to library services – completion of the task, time and effort, and reaction to the product or service. Similarly, but more specifically, Jeng (2006) proposed a usability model for academic digital libraries employing four operational usability criteria – effectiveness, efficiency, satisfaction, and learnability. Additionally, many other researchers conducted usability tests in DL environment (Eliassen et al., 1997; Battleson et al., 2001; Hammil, 2003; Blandford et al., 2004; Joo, Lin, & Lu, 2011).

In addition, researchers tried to evaluate digital libraries in different aspects. Shim and Kantor (1999) adopted Data Envelopment Analysis (DEA) to evaluate DLs, and proposed an evaluation framework that focused on two main dimensions of effectiveness and efficiency. Missingham (2003) introduced a unique concept of digital footprint representing the use of DLs through multidimensional measurement. Kim and Kim (2008) proposed 19 evaluation criteria that assess the quality of digital collections, and validated those criteria empirically based on the survey of users, librarians and administrators. Based on document analysis, Joo and Xie (2013) proposed a comprehensive pool of evaluation criteria in ten dimensions of academic digital libraries. Additionally, many other researchers suggested a variety of evaluation criteria such as suitability, accuracy,

costs, informativeness, timeliness, usefulness, and others (Kengeri et al., 1999; Kenney et al., 1998; Larsen, 2002).

2.6. Information Seeking Process Models

As this study focuses on the evaluation of search processes, representative previous models of information seeking processes are reviewed in this section. The author reviews several models that emphasize "process" in explaining information searches or retrieval.

The Ellis' model of information-seeking behavior is one of the most cited models focusing on search process. Ellis and his colleagues (Ellis, 1989; Ellis, Cox and Hall 1993; Ellis and Haugan 1997) characterized information seeking patterns from a series of empirical studies involving scientists, engineers, and social scientists in both academic and industrial settings. Based on empirical findings they have identified eight features of information seeking behavior as follows (Ellis and Haugan 1997):

- Starting (the search process): activities such as the initial search for an overview of the literature or locating key people working in the field;
- Chaining (moving from seeking): following footnotes and citations in known material or forward chaining from known items through citation indexes, or proceeding in personal networks;
- Browsing: variably directed and structured scanning of primary and secondary sources;
- Differentiating (to filter information): using known differences in information sources as a way of filtering the amount of information obtained;

- Monitoring (to keep up to date): regularly following developments in a field through particular formal and informal channels and sources;
- Extracting (relevant material): selectively identifying relevant material in an information source;
- Verifying (checking accuracy): checking the accuracy of information;
- Ending: activities actually finishing the information seeking process.

The strength of Ellis' model is that it was tested in a series of empirical studies and showed users applying multiple information seeking strategies during the search process. This model opened an avenue for researchers to pay attention to “process” when exploring information seeking behaviors.

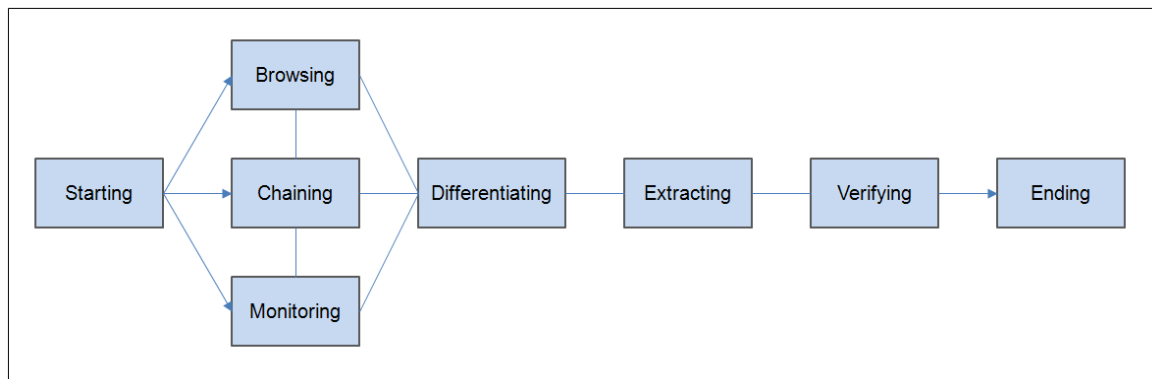


Figure 2-7. A stage process model of Ellis' information behavioral framework (Wilson, 1999, p.255)

Bates' (1989) “Berrypicking” is another earlier model that explores the patterns of information seeking behavior while emphasizing “search process.” In her model, a search process is evolving as a user searches for information a bit at a time using multiple search strategies. Each piece of information gathered leads to new ideas and consequently new queries can be formulated. This model highlights the iterative nature of users' search

process. She also identified six types of search strategies, including footnote chasing, citation searching, journal run, area scanning, subject searches, and author searching.

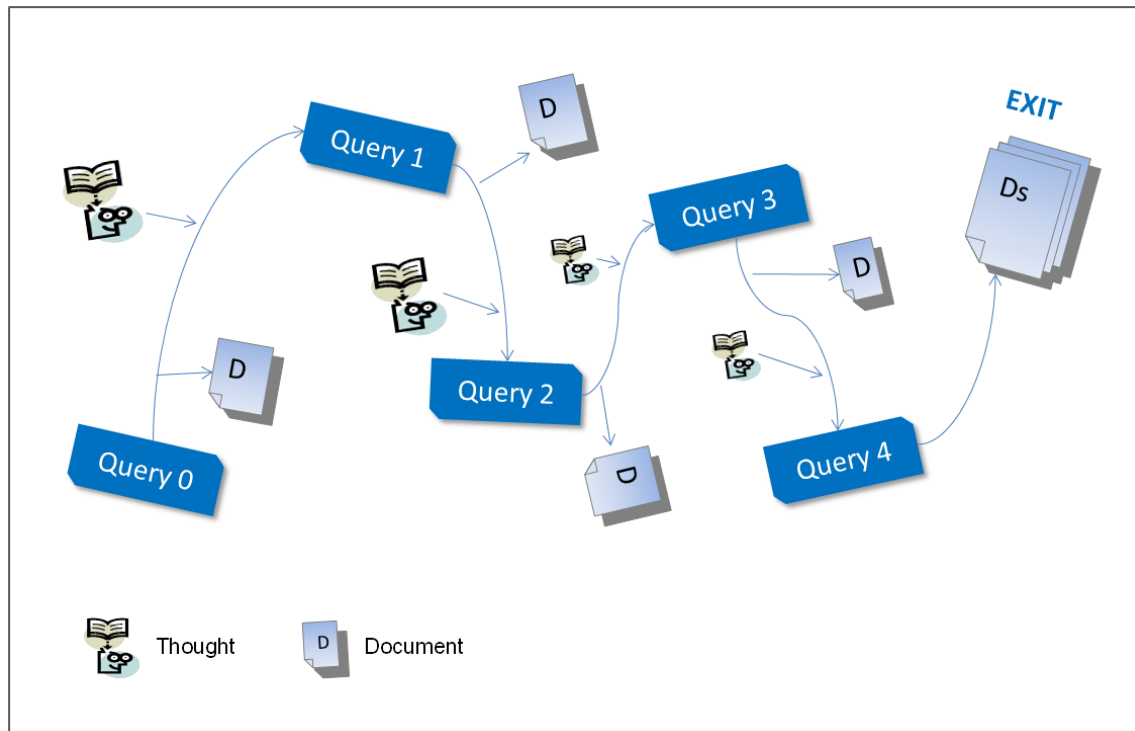


Figure 2-8. A berry-picking, evolving search (Bates, 1989, p.410).

Kuhlthau's (1991, 1993) Information Seeking Process (ISP) model integrates a search process with three different realms of human experiences: affective (feelings), cognitive (thoughts), and physical (actions). Kuhlthau's ISP model incorporates users' cognitive and affective experiences in understanding information seeking processes. The model describes the changes of thoughts and feelings of users along the stages of the process. The central assumption of the model is that a user's information seeking problem can be explained by uncertainty and confusion, which may lead to anxiety. Information seeking is, then, viewed as a process of construction in which a user progresses from uncertainty

to understanding as they seek for information. The search process is described in six stages of successive information searches: task initiation, topic selection, pre-focus exploration, focus formulation, information collection, and search closure.

- Initiation: becoming aware of the need for information (feeling of uncertainty)
- Selection: identifying general topic (feeling of optimistic)
- Exploration: investigating information on general topic (feeling of confusion and uncertainty)
- Formulation: formulating focus (uncertainty reduced, confidence increases)
- Collection: pertaining to focus (increased confidence and interest)
- Presentation: compiling information search and writing and presentation (relief, satisfaction/dissatisfaction)

Belkin et al. (1993; 1996) created a multi-faceted classification scheme of search strategies to characterize information-seeking behaviors. After empirically analyzing information seeking behaviors, they came up with a classification scheme on the basis of four behavioral dimensions:

- Method – whether a user is either searching for a particular information object, or scanning a set of information objects
- Goal – whether a user is learning about something or selecting something
- Mode – defines where a user is recognizing and specifying something
- Resource – whether a user is looking for information items or metadata about an information item

This multiple facets are used to represent a space of possible information-seeking strategies during an information seeking episode. Based on combination of the four facets, they classified sixteen distinct information-seeking strategies as shown in Table 2-4.

Table 2-4. Faceted classification of information-seeking strategies (Belkin et al. 1993, p.326)

ISS	Method		Goal		Mode		Resource	
	search	scan	learn	Select	recognize	specify	information	meta-information
1	✓		✓		✓		✓	
2	✓		✓		✓			✓
3	✓		✓			✓	✓	
4	✓		✓			✓		✓
5	✓			✓	✓		✓	
6	✓			✓	✓			✓
7	✓			✓		✓	✓	
8	✓			✓		✓		✓
9		✓	✓		✓		✓	
10		✓	✓		✓			✓
11		✓	✓			✓	✓	
12		✓	✓			✓		✓
13		✓		✓	✓		✓	
14		✓		✓	✓			✓
15		✓		✓		✓	✓	
16		✓		✓		✓		✓

Belkin and his colleagues (1995, 1996) suggested the information-seeking episode model by presenting a new lens to understand the flow of interactions between users and systems. In their research approach, the central process of information retrieval is user interaction with text, and accordingly the user is the central component of the IR system. Figure 2-9 illustrates Belkin's episode model that represents IR support for multiple interactions with information. In this model, an information seeking episode is viewed as a sequence of users' multiple interactions, and the type of interaction is dependent on the searcher's goal, problem, intention, situation, and others at each sequence. The strength

of the episode model is that it presents each interaction supported by a variety of processes, such as representation, comparison, visualization, and so forth. Each process can be initiated by some optimum combination of techniques for effective support of interaction. This model is unique in that it viewed the objective of an IR system as supporting multiple information-seeking behaviors, while placing users' interactions with text in the central process during the IR episode.

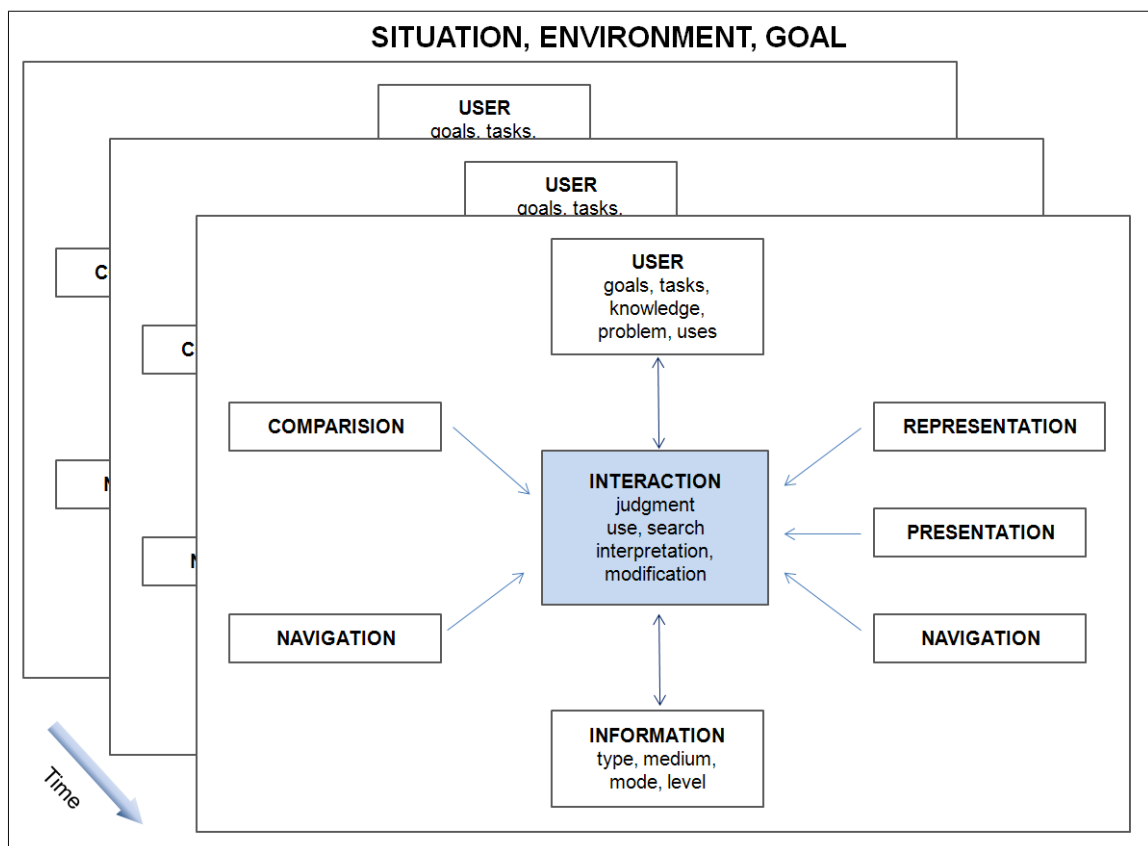


Figure 2-9. Belkin's episode model of interaction with texts: a model of IR support for multiple interactions (Belkin, 1996, p. 29)

Vakkari (2000a, 2000b, 2001) and Vakkari & Hakala (2000) investigated how the task performance process is connected to IR by examining graduate students' information

seeking processes in writing a thesis proposal. In particular, his research concern lay on how the search stages were related to the information types searched, to the search tactics and term choices, and to the level of judgments of relevance and full-texts. Based on a series of longitudinal studies, Vakkari (2001) has proposed a model of the information search process in task performance. This model summarizes the results of a series of studies and concludes that the stage of the task performance process would determine the information sought, search tactics applied, term choices selected, relevance judgments assessed, and documents obtained and used. The strength of this model is that it integrates and extends Kuhlthau's ISP into the task performance process model. Vakkari (2001) further refined and enhanced the major concepts in Kuhlthau's ISP in terms of search tactics, search terms, relevance feedback, and others. Ingwersen and Järvelin (2005, p. 199) addressed three strengths of Vakkari's model: (1) it offers clear-cut distinction between domain knowledge associated to work task performance and information seeking and retrieval knowledge; (2) work task stages and use of information are separated from search task execution; and (3) the introduction of a concept of "expected contribution," which refers to the experience gained by the actor, is also the contribution of the model. Moreover, his model has been validated in a series of empirical studies.

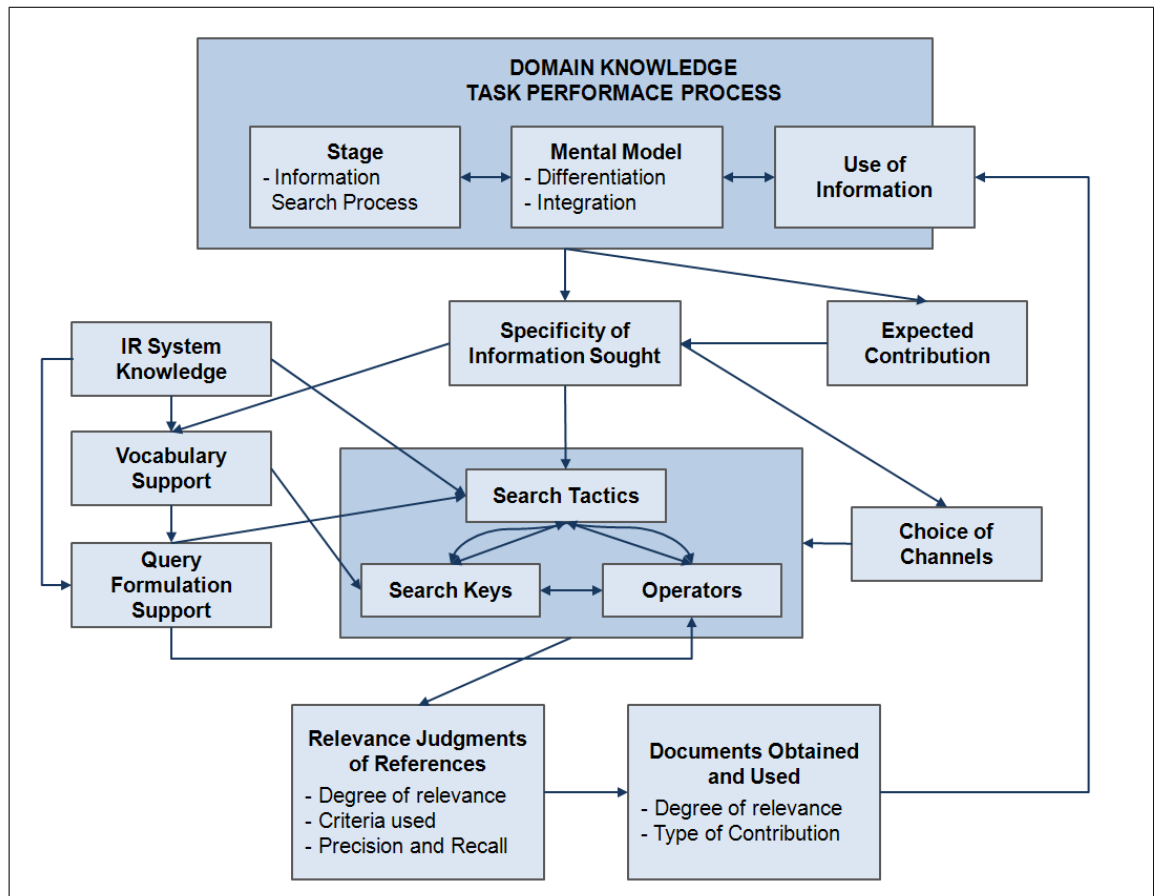


Figure 2-10. Vakkari's model of the information search process in task performance (Vakkari, 2001, p. 308)

Xie (2008a) created the *planned-situational interactive IR model* by consolidating macro- and micro-levels of interactions (Figure 2-11). She applied the planned model and the theory of situated action into interactive IR under the assumption that the nature of interactive IR is codetermined by a user's plan and confronting situation. Her model offers in-depth illustration of the micro-level of user goals (e.g., interactive intentions, retrieval tactics, and shifts of intentions), while incorporating social-organizational context simultaneously. The model implies that the determination of search strategies and their transitions comes from the products of plans and situations. Also, it highlights tiered levels of user goals/tasks and their representations, and identifies multiple dimensions of

search strategies, which consist of 12 types of interactive intentions and 11 types of associated tactics. The uniqueness of the model lies in that it investigates how users shift their current search goals, interactive intentions, and search tactics during the search process.

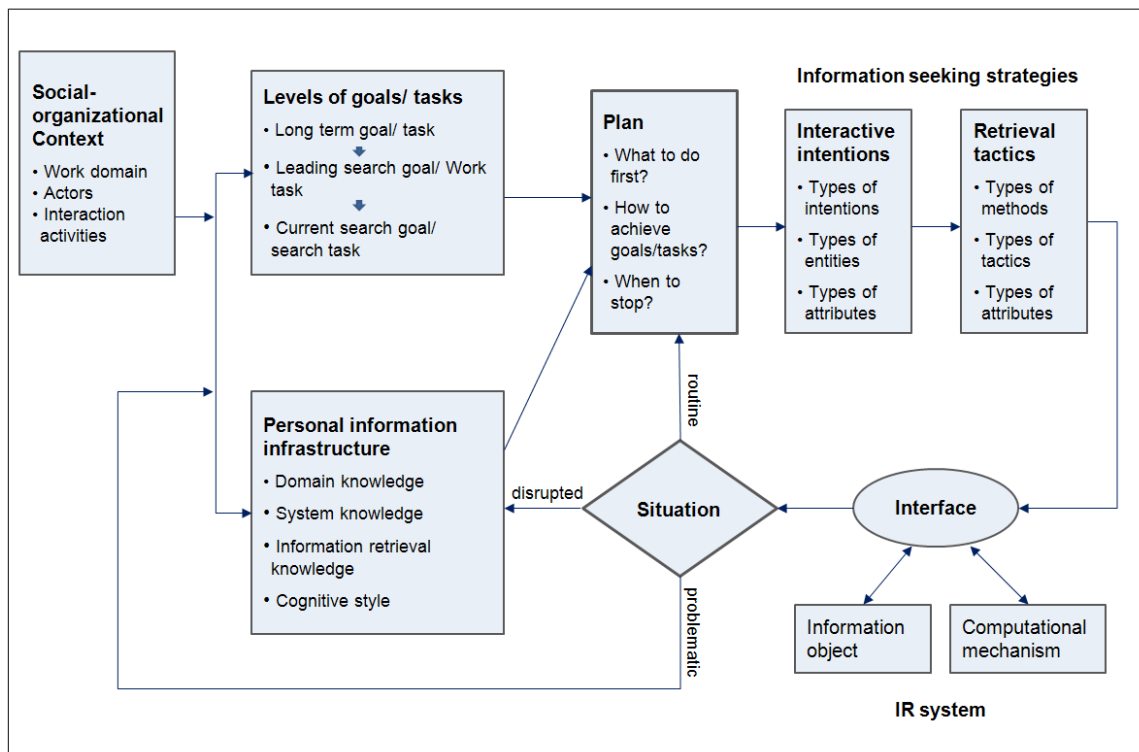


Figure 2-11. Planned-situational interactive IR model (Xie, 2008a, p. 216)

In addition, Xie (2009) examined how searchers plan their searches at three sequential stages of the search process: pre-focus, formation, and post-focus stages. At the pre-focus stage, searchers plan to explore different topics and to find general information in order to convert their information need from visceral to compromised need. At the formation stage, searchers' planning is focused on a specific topic as well as specific IR systems for searching. At the post-focus stage, searchers' plans are to develop more specific queries

and search tactics. In this ways, searchers' planning at different stages influences search topics, information used, and applied search strategies.

Xie (2008a) and her colleague (Xie & Cool, 2000) explored different types of search tactics and strategies, and investigated the shifts and sequential patterns among them. Xie and Cool (2000) identified six sub-tasks in which users would engage during the IR process to accomplish their search tasks in searching online databases: databases selection, query formulation, query reformulation, help mechanism access, results organization and display, and results delivery. Xie (2008a) further elaborated sub-tasks in search processes and came up with eleven types of user intentions that users need to exert to accomplish their search tasks: selecting databases, formulating search statement, re-formulating search statement, exploring, learning, monitoring, organizing, accessing, evaluating search results, keeping records, and using search results.

2.7. Limitations of Previous Research

Interactive IR has become one of the major areas in the discipline of information science. In particular, TREC Interactive Track laid the groundwork for the development of interactive IR methods and experimental designs that involve users in evaluation studies (Dumais & Belkin, 2005). Process-oriented information seeking models have contributed to the understanding of user search behaviors in search tasks. In addition, many researchers and practitioners have substantially exerted their efforts to digital library evaluation, and produced evaluation models specialized to digital libraries. However, there are some limitations of existing research in terms of evaluating user engagement

and system support in the context of digital libraries. This study addresses three limitations of previous literature in relation to interactive IR studies focusing on search process in digital library environments.

First, there has been little research that utilizes information-seeking process models to assess information search processes. As reviewed in this chapter, researchers generated various information-seeking process models, such as Ellis' model, Marchionini's model, Kulhthau's ISP, among others. These models have been useful to understand complex, dynamic nature of information seeking behavior during the search process. However, these models are created to conceptually understand information-seeking processes, rather than to practically assess user engagement and the quality of search processes.

Second, there are fewer interactive IR studies in the area of digital libraries, compared to other IR systems such as search engines and online databases. As reviewed in Section 2.5.4, major concerns in digital library evaluation have been usage, services, interface design and usability, and technologies. Less research has been done in relation to interactive IR evaluation in the area of digital library research. Most of digital library systems typically support both searching and browsing, as they are equipped with both search functions and topic categories for browsing (Zhang et al., 2011). Also, many digital libraries provide other various system features to help users implicitly and explicitly, and users interact dynamically with different system features while using digital libraries (Zhang et al., 2011; Huang & Xie, 2011; Miller et al., 2012). Therefore, for digital libraries, it would be necessary to assess various types of interactions between users and system features based on interactive IR method.

Third, less research focused on “search process” in interactive IR evaluation. Interactive IR, represented by TREC Interactive Track, contributed to the inclusion of users in the design of IR evaluation studies. In interactive IR, however, IR evaluation has been conducted based on output variables, which measure the final product of search process. There are few studies that assessed user-system interaction processes rather than search results. Interactive IR still concerns more on performance-based evaluation, rather than process-based evaluation.

2.8. Conceptual Research Model of User-System Interactions in Digital Libraries

The limitations of previous literature call for the need to assess information search process based on the analysis of user search tactic application and associated system support in digital library environments. This study 1) explores users' search tactic application, 2) assesses search process in terms of system support, difficulty, and satisfaction in applying search tactics and 3) effects of search tactic selection and system support on search outputs. This section establishes a conceptual research model by extending Xie's (2003; 2008a) model of user engagement and system support. Also, the conceptual research model incorporates the set of search tactics suggested by Xie and Joo (2010). Figure 2-12 summarizes the conceptual research model of this dissertation. As shown in the diagram, this study assumes that a search process consists of user engagement and system support. User engagement can be represented by sequential application of search tactics, while system support can be provided by the forms of system features. This study analyzes users' search tactic behavior as a way to investigate

user engagement in an IR process. Also, this study assumes that user engagement and system support affect search outputs, which are the results of a search process.

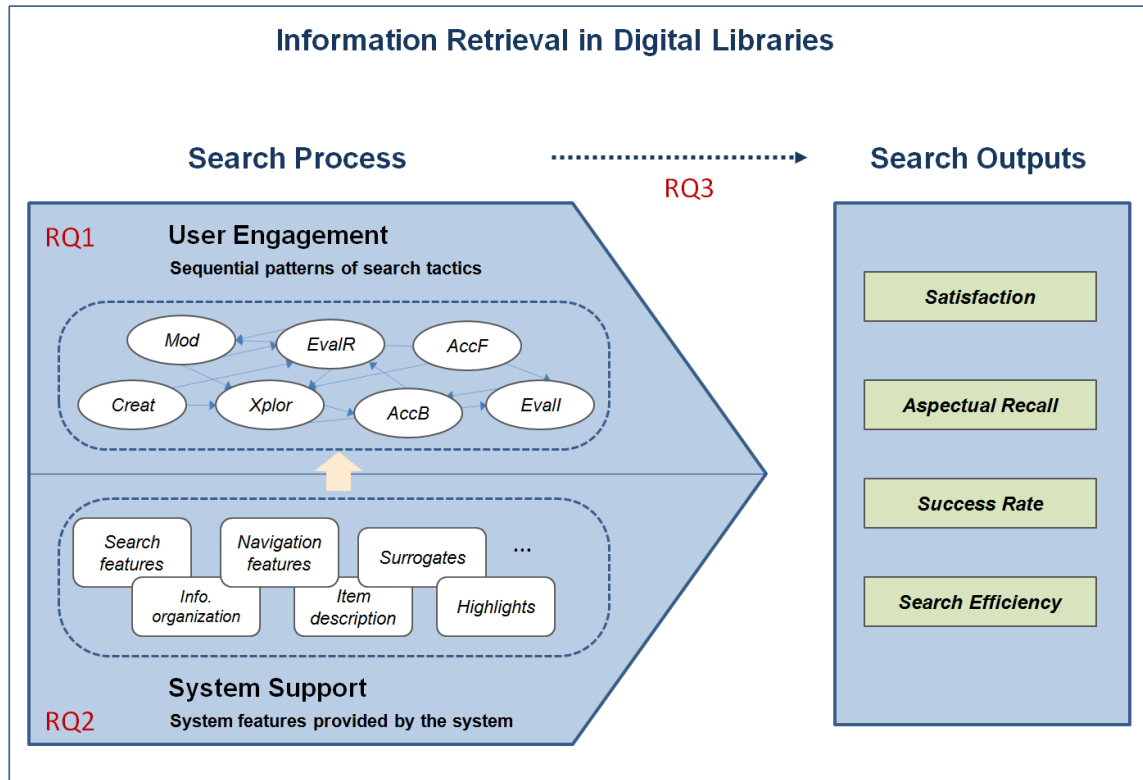


Figure 2-12. Conceptual research model

Xie's (2003; 2008a) model of system support and user engagement laid the fundamental basis for the research model of this study. In her model, she defines an IR process as a collaborative process of both user engagement and corresponding system support.

Interactive IR process is represented by balancing two different dimensions of labor, namely user engagement and system's reactive support in terms of: 1) to what extent a user should exert his/her effort to control for the IR process and 2) to what extent the system should support the user to easily proceed to the IR process (Xie & Cool, 2000; Xie, 2003).

This study assumes that user engagement is represented by users' application of search tactics. Therefore, in the conceptual research model, search tactics are the key concept in explaining user engagement. As reviewed in Section 2.3, researchers identified various types of search tactics. This study's conceptual model adopts Xie and Joo's (2010a) identification of search tactic types as shown in Table 2-6. This study basically investigates which search tactics are selected in accomplishing different types of search tasks in using digital libraries. Table 2-6 was used as a coding scheme for this study.

Table 2-5. Coding scheme of search tactics

Code*	Types of search tactics	Objective
Creat	Creating search statement	Come up with a search statement for searching.
Mod	Modifying search statement	Change a previous search statement to specify or broad search results.
EvalI	Evaluating individual item(s)	Assess relevance/usefulness of an item, or authority of an item.
EvalR	Evaluating search results	Quickly assess the relevance of search results.
Rec	Keeping a record	Keep records of metadata of an item(s) before accessing it/them.
AccF	Accessing forward	Go to a specific item or web page that has not been accessed in the search by using direct location, tracking meta-information, or hyperlinks.
AccB	Accessing backward	Go back to a previous page by using direct location, tracking meta-information, or hyperlinks
Lm	Learning	Gain knowledge of system features, system structure, domain knowledge, and database content.
Xplor	Exploring	Browse information/items in a specific information system.
Org	Organizing	Sort out a list of items with common characteristics.
Mon	Monitoring	Examine the search process or check the current status.
Obt	Obtaining	Obtain information in physical or electronic formats.

* Abbreviation for each type of search tactic

More importantly, this study focuses on patterns of search tactic selection in search processes. As shown in Figure 2-12, the conceptual research model strives to trace transitions of search tactics which are needed to complete a search task. Based on the conceptual research model, an empirical study examines 1) selection of search tactics, 2) changes of search tactic selection over time in a session, and 3) transitions in search tactics. This analysis on search tactics will be used to investigate user engagement in search process.

The other side of interest is the evaluation of search processes. To evaluate search process, different levels of user engagement and system support are measured for search tactics. At the physical level, the object of measurement is users' search tactic application. In this physical level, it is investigated how users put labor into the search process by manipulating different types of search tactics. Basically, frequency of and time spent on applying each type of search tactic are measured. Also, the author explores transitions and patterns of search tactic application. Researchers have identified a range of measures to represent users' information search behavior, such as time spent on a search session, numbers of pages viewed, and frequency of navigational tool use (Palmquist & Kim, 2000; Kim & Allen, 2002). This study split a search session into a series of sequential search tactics. That is, a search session consists of sequential transitions of search tactics. Accordingly, in this study, the unit of analysis is a search tactic that users apply during the search process.

At the cognitive level, two perceptual aspects are investigated -- perceived system support and difficulty. An IR process requires the searcher's multiple perceptions to proceed with the search process (Ingwersen, 1992). Interactive IR evaluation concerns searchers' perception caused by different events during the search session (Joho, 2009). In this conceptual research model, the author is interested in user perception of search tactic application in a search process. In interactive IR research, cognitive state has been estimated by different variables such as usefulness, cognitive load, difficulty, confidence, uncertainty, and others (Kuhlthau, 1993; Belkin et al., 2009; Lin 2002, 2005; Tenopir et al., 2008; Gwizdka, 2010). Among these variables, this study measures system support and difficulty in terms of search tactic application. Users' perceived system support

reflects their cognitive responses to the system support they experience during the search process. Difficulty is related to cognitive requirements imposed by the search system or the task itself. Difficulty is one of subordinate elements of the cognitive load in the IR context (Gwizdka, 2010).

At the affective level, this study attempts to measure users' perceived satisfaction level for the application of search tactics. Users' emotional state is an important factor in IR evaluation, as it could be leveraged to improve search results (Bennett et al, 2012).

Searchers' affective status has been measured by different emotional variables, such as satisfaction, pleasure, feelings of achievement, anxiety, annoyance and frustration (Baroudi et al., 1986; Saracevic, 1991; Lopatovska & Mokros, 2008; Tenopir et al., 2008).

Among these variables, user satisfaction is one of the widely selected evaluation criteria, as it is one of the main goals of IR systems (Cheng, Hu, & Heidorn, 2010). From the measurement of perceived satisfaction, this study tried to estimate to what extent users are satisfied with the process of search tactic application at the affective level.

2.9. Chapter Summary

This chapter defines key concepts and terms in this dissertation (Section 2.1). This study reviews previous discussion on user engagement and associated system support. Existing literature that identified different types of search tactics is reviewed as this study investigates search process at the search tactic level (Section 2.2 & 2.3). Then, there is a comprehensive review about IR system evaluation. In particular, interactive IR studies are reviewed focusing on TREC Interactive Track, evaluation criteria and effects of task

types (Section 2.4 & 2.5). Information seeking process models are also covered in this chapter (Section 2.6). Section 2.7 addresses the limitations of previous research in terms of process-driven evaluation and interactive IR studies in the area of digital libraries. Based on the literature review, a conceptual research model has been suggested in Section 2.8, which includes the relationships between search processes and search outputs.

CHAPTER 3. METHODOLOGY

The research purpose of this study is to investigate user engagement and system support during a search process in using digital libraries. The goal of this study is to explore users' search tactic application and to identify different types of system support for search tactics. In addition, the author examined the effects of search tactic patterns and system support on IR outputs. This chapter describe the research methods used to answer the research questions.

3.1. Sample

An empirical user study was designed to observe users' diverse interactions with digital libraries in different search task situations. The data were collected from sixty students from the University of Wisconsin-Milwaukee (UWM). University students are one of the major user groups of digital libraries. Undergraduate and graduate students utilize resources of digital libraries for their academic tasks, such as research projects, class assignments, or personal interests (Lee, Paik, & Joo, 2012). To better represent the population of UWM students, the author considered the proportions of sub-groups by major and academic status. Sixty subjects were recruited between May and September in 2012. Printed flyers, listserve emails, and referrals were employed to recruit participants across the campus. Table 3-1 presents the proportions of research subjects participated in this study. In the process of recruitment, all the subjects were asked whether they have basic-level computer literacy, including the skills of accessing the Internet and manipulating web browsers. Any student under 18 years old were excluded from the

study. As an incentive, \$30 were given to a subject for completion of research participation.

Table 3-1. Proportions of research subjects by major and academic status

	Humanities/ Arts	Social sciences	Sciences/ Engineering	Total
Graduate	10 (16.7%)	10 (16.7%)	10 (16.7%)	30 (50%)
Undergraduate	10 (16.7%)	10 (16.7%)	10 (16.7%)	30 (50%)
Total	20 (33.3%)	20 (33.3%)	20 (33.3%)	60 (100%)

Table 3-2 presents the demographic characteristics of the subjects participated in this study. More than 90% of the subjects were younger than 40 years old. They were all enrolled in either undergraduate or graduate programs at the time of research participation. With regard to computer literacy, the participants rated themselves at least intermediate level, which indicates they believe that they are fluent with using a computer for finding information from the Internet. Thus, all the subjects of this study satisfied the minimal requirement of computer literacy. More than half of the subjects never used LOC-DL before by accounting for 56.7%. There was no subject who used LOC-DL often or daily base. About 26.7% of the subjects answered "rarely use" while 16.7% were "occasionally use." All sixty subjects were native speakers of English.

Table 3-2. Demographic characteristics of subjects

	Category	Number	Percentage
Gender	Male	26	43.3%
	Female	34	56.7%
Age	18-29	43	71.7%
	30-39	12	20.0%
	40-49	4	6.7%
	50-59	1	1.7%
	60+	0	0%
Ethnicity	African American	1	1.7%
	Asian	1	1.7%
	Caucasian	44	73.3%
	Hispanic	7	11.7%
	Native American	1	1.7%
	Not Responded	6	10.0%
Computer skills (self-claimed level)	Expert	1	1.7%
	Advanced	24	40.0%
	Intermediate	35	58.3%
	Beginner	0	0.0%
Use frequency of LOC-DL	Use daily	0	0.0%
	Often use	0	0.0%
	Occasionally use	10	16.7%
	Rarely use	16	26.7%
	Never use	34	56.7%

3.2. Data Collection Methods and Procedures

3.2.1. Data Collection Methods

Data were collected from multiple methods: pre-questionnaires, screen recording, transaction logs, think-aloud verbal protocols, and post-questionnaires.

First, pre-questionnaires were used to collect demographic information about participants, such as gender, age, race, major, computer skills, and others.

Second, users were asked to conduct three search tasks using Library of Congress Digital Libraries (LOC-DL). All of the users' search activities were recorded using Morae usability testing software (<http://techsmith.com/morae.html>) to collect users' behavioral data, such as pages viewed, input device operations (e.g., clicks, keystrokes, etc.), facial expressions, and voices.

Third, subjects were asked to verbalize their intentions, thoughts and feelings in relation to their search activities during the search process. All think-aloud utterances along with facial expression and voice tone were recorded using Morae software.

Fourth, post-search questionnaires were used to measure subjects' perceptions of system support, difficulty, and satisfaction about search tactic application after the search. Figure 3-1 summarizes the data collection procedures of this study.

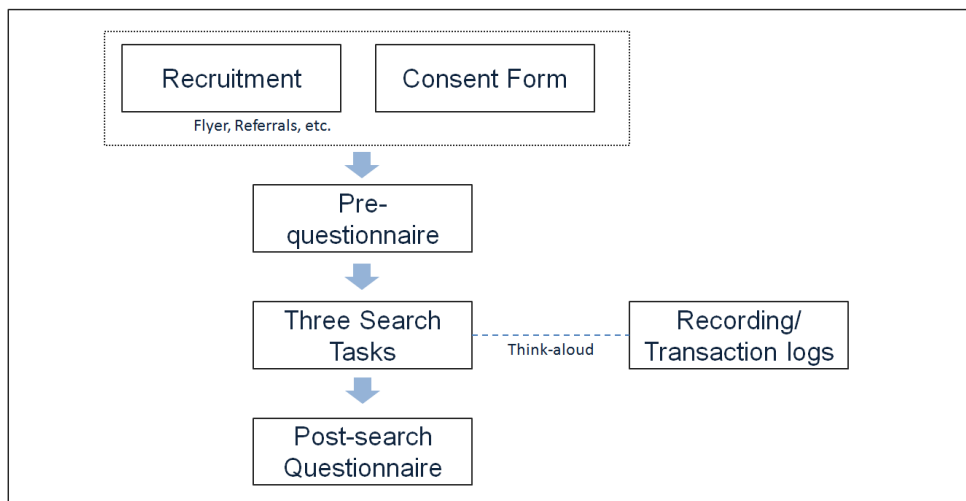


Figure 3-1. Data collection procedures

3.2.2. Selected Digital Library System

For this user study, Library of Congress Digital Libraries (LOC-DL) was selected as a test digital library system. LOC-DL is one of the representative national-level digital libraries in the United States, which covers a wide variety of topics. By selecting a currently operating digital library system, instead of an experimental system, the author aimed to observe users' interactions with real resources of a real system that reflect more real situations. In addition, digital libraries create a new searching environment in which many of the searchers are novice users (Xie 2009), so this study intends to uncover users' unique search behaviors in digital libraries, which have not been widely studied in IR research. To be more specific, LOC-DL was chosen as an IR system to be examined based on the following reasons:

- *Coverage of topics* — LOC-DL covers a wide variety of topics such as history, maps & geography, biography, arts & culture, religion, and philosophy, amongst others.
- *Resource formats* — LOC-DL offers multiple formats of sources such as text, images, audio files, video clips, and maps.
- *Search strategies* — LOC-DL provide a variety of search features in support of different types of search tactics.
- *Help features* — LOC-DL offers different types of explicit help features, including help pages, FAQ, search finding aids, instructional pages, etc.
- *Credibility of contents* — resources of LOC-DL are originated from reliable, trustworthy sources or entities.

- *Representation of digital libraries in academia* — LOC DL is one of representative national-level digital libraries run by Library of Congress.

This study involves digital library systems, so the results of this study cannot be generalized to other IR system settings. External validity indicates how well the results of a study can be generalized across different populations and settings. In this study, all the subjects were recruited from different disciplines including humanities and arts, social sciences, sciences/ engineering, and they represents both undergraduate and graduate student groups. Therefore, the results of this study can be generalized into the setting of digital library uses in a research university.

3.2.3. Search Tasks

In this study, three types of search tasks were designed to explore users' interactions with LOC-DL, including known-item search task, specific information search task, and exploratory search task. Search task types can be classified by search results that a user intends to obtain. *Known-item searching* refers to finding an item when a user knows particular information about that item, such as author, title and so forth. *Specific information searching* represents looking for exact data or a fact. *Exploratory searching* indicates looking for items with common characteristics (Xie, 2008a). Using multiple types of search tasks, the author planned to observe more diverse user engagement and corresponding system support during the search process.

First, as a known-item search task, subjects were asked to find a video clip of "Coca-Cola advertisements in 1964." LOC-DL has special collections about "Fifty years of Coca-

Cola." In the collections, there are several video clips of Coca-Cola advertisements, and subjects were asked to locate one of them broadcasted in 1964.

Second, as a specific information search task, subjects were asked to find who were the four US presidents assassinated and when they were assassinated. This task requested users to find very specific factual pieces of information. In this task, they were asked to find the names of four president who were assassinated while in office (Lincoln, Garfield, McKinley, Kennedy) and the dates of the assassinations (1865, 1881, 1901, 1963).

Therefore, in total, eight pieces of specific information were supposed to be searched to successfully complete this search task. LOC-DL has a special collection on American presidents, which includes information about president assassinations in the United States.

Third, as an exploratory search task, subjects were asked to collect as many aspects as possible on a certain topic within eight minutes. The selected topic was "Jackie Robinson's life and his career as a major league baseball player." LOC-DL has a special collection on Jackie Robinson with various aspects of information ranging from overview, timeline, essays, photos, achievement, and to his family. Subjects were allowed to apply any search strategies they wanted to solve the search task within the boundary of LOC-DL. To objectively calculate aspectual recall rates, subjects were instructed either to copy and paste the findings to the MS-Word file or to speak out whether to use the information they accessed. Table 3-3 summarizes three types of search tasks designed in this study. For simplicity sake, task ID numbers are used to indicate each type of search task throughout this dissertation.

Table 3-3. Three search task types: known-item search (Task 1), specific information search (Task 2), and exploratory search (Task 3)

Task ID	Search task type	Task	Time limit
Task 1	Known-item search	1. Find a Coca-Cola advertisement video clip in 1964	5 minutes
Task 2	Specific information search	2. Who are the four US presidents assassinated during their presidency? In which year was each of the president assassinated?	5 minutes
Task 3	Exploratory search	3. Assume that you are supposed to write a final report on Jackie Robinson's life and his career as a major league player. Please collect as many aspects as possible that could be useful for your report (e.g., biography, achievement, images, teams, records, etc.)	8 minutes

3.3. Data Analysis

Multiple data analysis techniques were selected to answer the proposed research questions. To answer RQ 1, exploratory analysis methods were used to quantitatively model users' search tactic patterns in three different search tasks. Transaction logs were coded into types of search tactics using the coding scheme identified in Table 2-6 (see Chapter 2). Ten percent of the total sessions (18 sessions) were coded by two coders to ensure the coding reliability. The first coder initially coded search tactics by analyzing both transaction logs and video records, and the second coder checked whether the first coder's coding was adequate. The inter-coder reliability turned out 95.2%, which was calculated as the ratio of the number of agreed coding decisions over the total number of

the initial coding decisions. Table 3-4 summarizes the data analysis strategies used in this study for each research question.

Table 3-4. Data analysis methods

Research Question	Data Collection	Data Analysis
1. Users search tactic selections and patterns	a) Transaction logs b) Screen recording c) Think-aloud protocols	a) Descriptive statistics b) Comparison by task type (repeated-measures ANOVA) c) Search tactic changes (kernel regression) d) Transitions in search tactics (Markov switching analysis) e) Session clustering (hierarchical clustering and multi-dimensional scaling)
2. Types of system supports; Perceived system support, difficulty, and satisfaction	a) Transaction logs b) Screen recording c) Think-aloud protocols d) Post-questionnaires	a) Open coding b) Descriptive statistics c) Correlation analysis (Pearson r) d) ANOVA
3. Effects of user search tactic selections and system supports on search outputs	a) Results of RQ1 b) Results of RQ2 c) Post- questionnaires	a) Correlation analysis (Pearson r) b) Multiple regression c) Structural equation modeling

For RQ 1, descriptive statistics was basically used to count frequency of and time spent on search tactics in three different search tasks. This descriptive analysis identified the most frequently used search tactics in searching digital libraries. Time data further detailed search tactic selections in digital library searches. In addition, the author examined the effect of task type on search tactic frequency and time based on ANOVA.

In addition, the author traced how users' search tactic selection changed within a session over time. As patterns of search tactic changes exhibited non-linearity, kernel regression was applied using MATLAB (<http://www.mathworks.com/products/matlab/>). Kernel regression is a nonparametric method in statistics that estimates non-linear patterns among variables (Eubank 1999; Takeda et al., 2007). The kernel regression can be represented with a regression function, $g(x_i)$:

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

(where $g(x_i)$ = unknown regression function; ε_i = the independent and identically distributed zero mean errors). $g(x_i)$ is estimated by the following nonparametric function:

$$\hat{g}(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)}$$

(where K = kernel function; h = bandwidth). The requirements for the kernel function are:

$$\int K(z)dz = 1 ; \quad \int zK(z)dz = 0 ; \quad \int z^2K(z)dz = c$$

where c is a constant value (Takeda et al., 2007). Epanechnikov kernel and Silverman's rule of bandwidth calculation were selected in this analysis (Silverman 1986; Takeda et al., 2007):

$$K(u) = \frac{3}{4}(1 - u^2)\mathbf{1}_{\{|u| \leq 1\}}$$

$$h = 1.06 * \sigma * n^{-\frac{1}{5}}$$

Epanechnikov kernel and Silverman's rule are commonly used to estimate nonlinearity between two random variables. Based on kernel regression, this study attempted to model the patterns of search tactic changes over time during a session. Each length session was standardized between 0 (starting point) and 1 (ending point), and tactic occurrence probabilities were calculated for each 0.001 point. The calculated tactic occurrence probabilities at 1,000 points were estimated based on kernel regression. R-square and Root Mean Square Error (RMSE) were selected to indicate a model fit. Figure 3-2 shows an example of kernel modeling of Xplor tactic changes in Task 3. The y-axis indicates the probability of search tactic occurrence while the x-axis indicates time flow in a session.

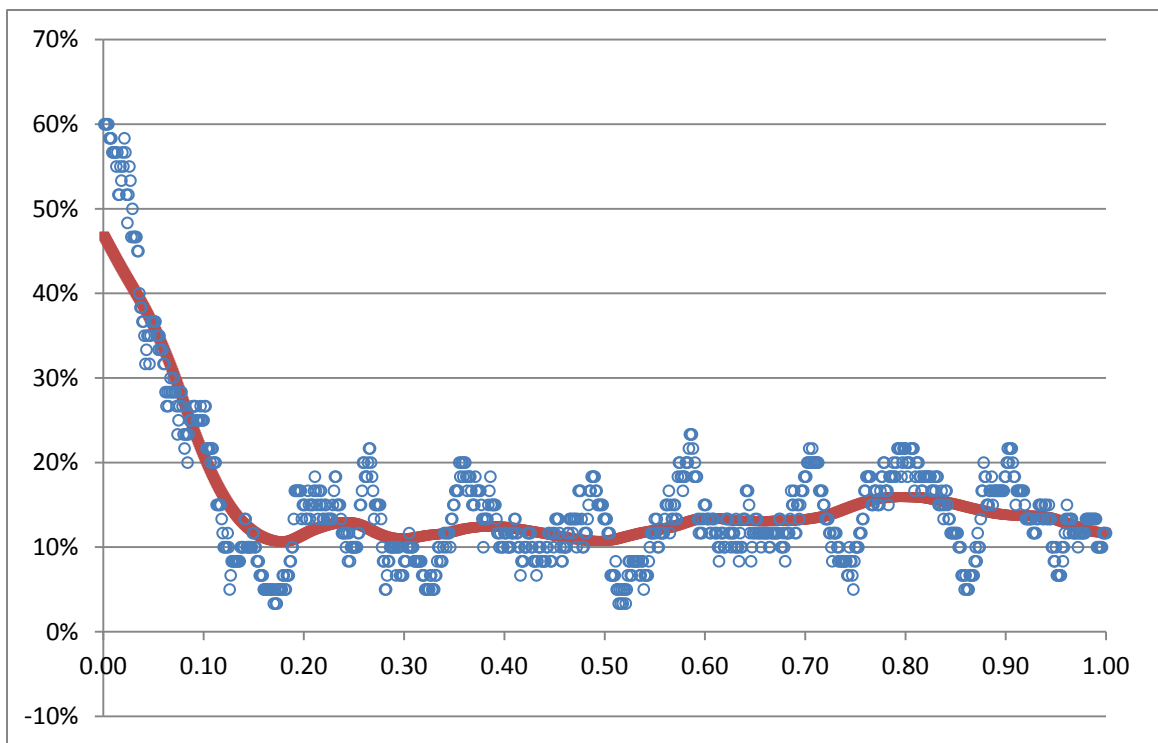


Figure 3-2. Example of kernel regression modeling (circle - observation; line - regression estimation; y-axis = probability of tactic application; x-axis = time flow in a session)

Furthermore, this study tries to identify search strategies applied by analyzing transitions between search tactics. A regular Markov switching chain was used to model search tactic transitions during the search process. A Markov chain is a mathematical method to calculate the probability of transitions between one state and another (Chen & Cooper, 2002). A sequence of search tactic applications was traced to identify which paths of search tactics are frequently used to complete search tasks. By modeling the transition and associated probabilities, we can predict what an expected user is likely to do in the search process (Guo et al., 2008). Additionally, observed sessions were analyzed by their characteristics. Based on hierarchical clustering and multidimensional scaling (MDS), search sessions were classified into different groups, such as result evaluation oriented or browsing oriented sessions. As a clustering method, Ward's minimum variance method was applied (Ward 1963) while Minkowski measures were used for distance. A dendrogram was used to interpret the clustering result. For MDS analysis, a dissimilarity matrix was created based on Euclidean distance, and all observed sessions were projected in a two-dimensional space for interpretation.

In answering RQ2, both qualitative and quantitative analyses were used. Firstly, types of system support were identified from observations of user-system interactions. Open coding was used, which is the process of breaking down, examining, comparing, conceptualizing, and categorizing unstructured textual transcripts (Glaser & Strauss, 1967; Strauss & Corbin, 1990). Secondly, system support was measured for each type of search tactic using a post-search questionnaire. The post-search questionnaire was designed to assess users' perceptions of system support for applying each type of search tactic. Post-questionnaires were administered only to the sessions of exploratory search tasks (Task

3). From the pilot test, the author found that it took too long to gather answers from all three tasks, and the participants of the pilot test showed fatigue in answering too many questions. Therefore, the post-search questionnaire was conducted only for Task 3. A five-point Likert scale was employed to numerically represent levels of system support. Subjects were asked to rate their perception to what extent they were supported from LOC-DL while applying each type of search tactic. Additionally, perceived difficulty was measured for each type of search tactic based on a five-point scale. The author was interested in how system support and difficulty levels would be related to each other. Correlation analysis was conducted to look into the relationship between users' perceived system support and perceived difficulty in applying search tactics during the search process. At the affective level, user satisfaction was of interest in this study. Again, five-point scale was used to gauge user satisfaction level with the process of applying different types of search tactics.

For the results of RQ 3, this study examined the effects of users' search tactic selections on IR outputs based on multiple regression. Required sample size was calculated for multiple regression based on Cohen's effect size (Cohen, 1998). To achieve 0.8 statistical power level with an anticipated effect size (f^2) of 0.30, fifty six observations are needed in a regression model of seven predictors at an alpha level of 0.05. This study satisfied the minimum number of sample size for a multiple regression with seven independent variables. Collinearity was diagnosed based on Tolerance and VIF (variance inflation factors) indices. A tolerance of less than 0.20 and a VIF of 5 and above were used as the criterion of a significant collinearity problem (O'Brien, 2007). If significant collinearity was detected, a stepwise method was used to avoid multicollinearity between variables.

In Task 1, the author analyzed the causal relationship between search tactic selections and session efficiency. Since session lengths varied by session, the frequency of search tactics was standardized by time. "Average number of search tactics per minute" was used as an independent variable in the analysis, which indicates standardized search tactic frequency by time. Seven independent variables were used: frequencies of Creat, Mod, AccF, AccB, EvalR, EvalI, and Xplor respectively. Less frequently observed search tactics were excluded from the regression model, such as Mon, Lrn, and Org. Collinearity was first checked, and if it was detected, a stepwise method was selected as an alternative way. Session length was selected as a dependent variable. Session length has been often selected as a measure that represents session efficiency (Shackel, 1991; Battleson et al, 2001; Joo, 2011). A shorter session is considered more efficient, which reveals that users can complete a search task quickly with less effort. Figure 3-3 presents a multiple regression model that examines the effect of search tactic selection on search efficiency. Seven hypotheses were established in this analysis. The author assumed that the standardized frequency of each type of search tactic would influence search efficiency measured by session length.

(H1: 1-7) Frequency of [Creat, Mod, EvalR, EvalI, AccF, AccB, Xplor] tactics affects session length.

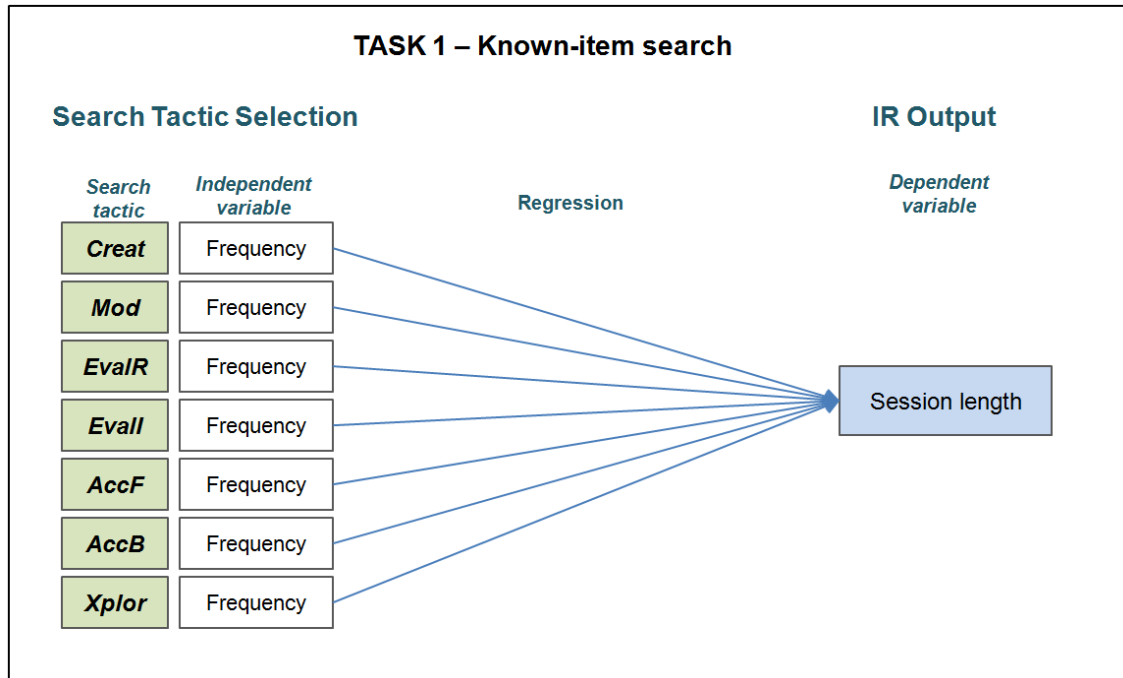


Figure 3-3. Examining the effect of search tactic selection on session efficiency (Task 1)

For Task 2, the effect of search tactic selection on success rate were investigated. Task 2 requested subjects to find eight pieces of factual information. As a measure for search output, success rate was adopted. Success rate refers to the percentage of requested items that users found during a session. The number of information pieces a subject found out of the eight requested pieces was measured. As independent variables, both frequency of and time spent on each type of search tactic were selected. Again, since significant collinearity existed among independent variables, a stepwise predictor entering method was used in the regression analysis. Figure 3-4 describes a multiple regression model to examine the effect of search tactic selection on success rate. Twelve hypotheses are established for the causal relationship analysis in Task 2.

(H2: 1-7) Frequency of [Creat, Mod, EvalR, EvalI, AccF, AccB, Xplor] tactics affects success rate.

(H2: 8-12) Time spent on applying [Creat, Mod, EvalR, EvalI, Xplor] tactics affects success rate.

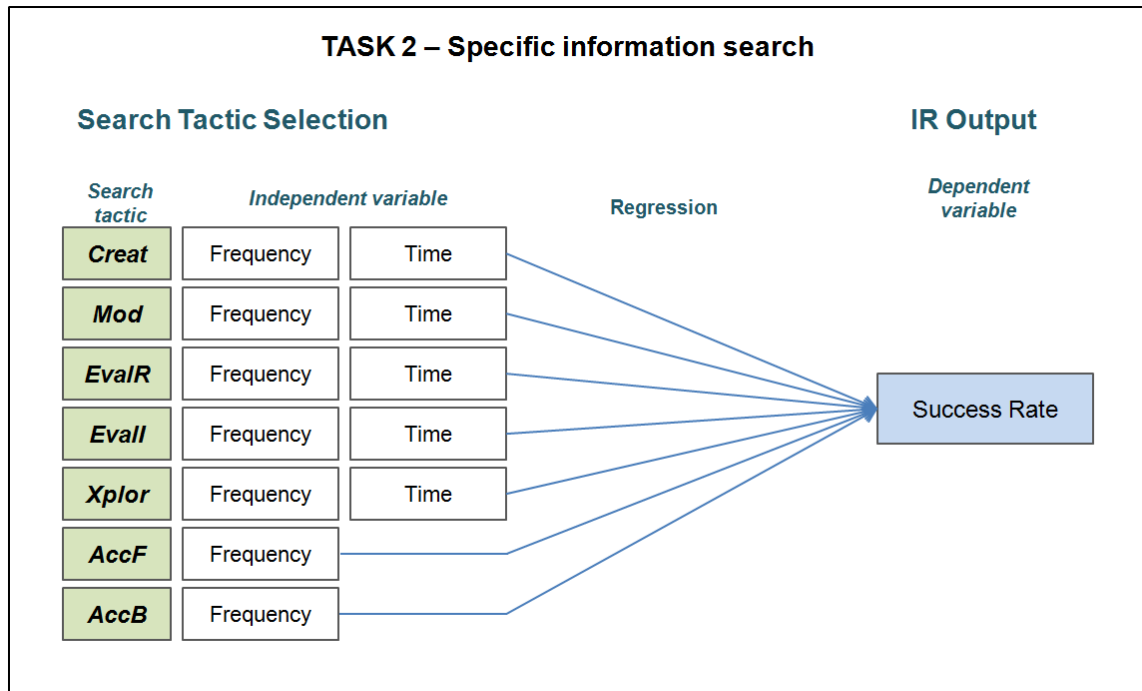


Figure 3-4. Examining the effect of search tactic selections on success rate (Task 2)

For Task 3, "aspectual recall" and "satisfaction with search results" were selected to represent the quality of search outputs. To objectively judge which items were finally selected for the task, subjects were instructed to either save findings to an MS-Word file or verbalize their intention to use an item (e.g., "I'm going to take the image (S40)."). In this way, the coders were able to objectively determine which items were chosen by subjects while conducting the Task 3. Based on a pooling method suggested by Yuan and Belkin (2007), aspectual recall was determined by pooling all of the aspects identified for

the task by all of the subjects. The other output measure was satisfaction with search results. The author tried to measure the quality of search results at the affective level. The post-search questionnaires directly asked subjects to rate to what extent they were satisfied with obtained information using a five-point Likert scale.

RQ 3 tried to investigate how users' search tactic selection and system support would influence search outputs. After the quality of search outputs were measured by aspectual recall and satisfaction level, the author examined the causal relationships 1) between users' search tactic application and search outputs and 2) between users' perceived system support and search outputs. Since there were two output variables (dependent variables) and two independent variables, four separate multiple regressions were conducted in this study.

Figure 3-5 presents four multiple regression models to investigate the effect of search tactic selection on IR outputs. The author assumed that users' search tactic selections would influence search outputs in Task 3, and the following twenty four hypotheses are established:

(H3: 1-7) Frequency of [Creat, Mod, EvalR, EvalI, AccF, AccB, Xplor] tactics affects aspectual recall.

(H3: 8-14) Frequency of [Creat, Mod, EvalR, EvalI, AccF, AccB, Xplor] tactics affects satisfaction with search results.

(H3: 15-19) Time spent on applying [Creat, Mod, EvalR, EvalI, Xplor] tactics affects aspectual recall.

(H3: 20-24) Time spent on applying [Creat, Mod, EvalR, EvalI, Xplor] tactics affects satisfaction with search results.

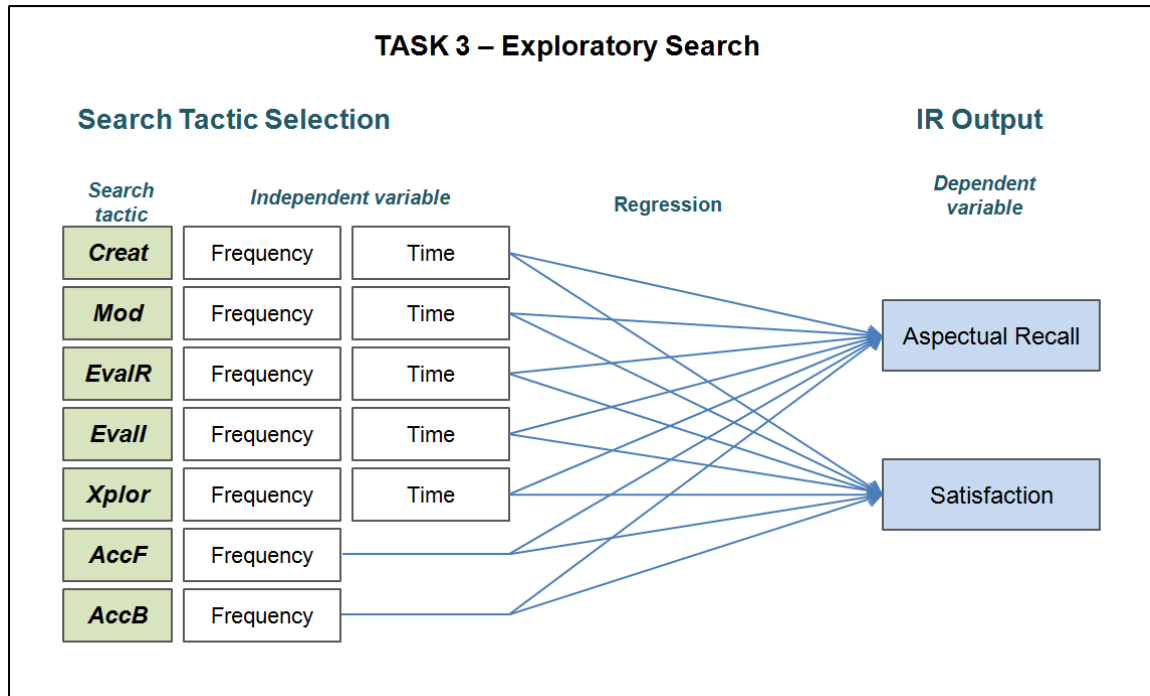


Figure 3-5. Examining the effect of search tactic selections on IR outputs (Task 3)

Figure 3-6 delineates another two multiple regression models that investigate the relationships between perceived system support and IR outputs. In this regression model, it was assumed that perceived system support for search tactics would influence aspectual recall and satisfaction with search results. The following fourteen hypotheses were identified for statistical test.

(H3: 25-31) System support for [Creat, Mod, EvalR, EvalI, AccF, AccB, Xplor] tactics affects aspectual recall.

(H3: 32-38) System support for [Creat, Mod, EvalR, EvalI, AccF, AccB, Xplor] tactics affects satisfaction with search results.

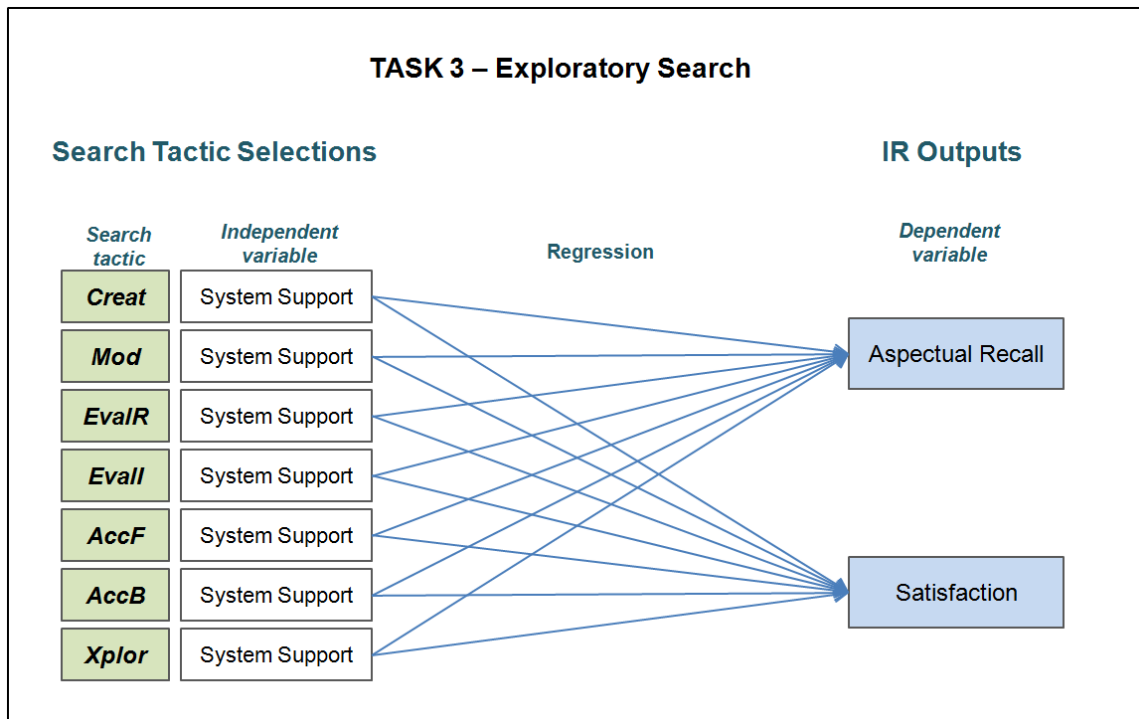


Figure 3-6. Examining the effect of system support on IR outputs (Task 3)

Additionally, structural equation modeling (SEM) was employed to comprehensively analyze the relationships among process-related variables and output variables. SEM, which is also known as the covariance structural model is a multivariate statistical analysis technique for establishing, estimating, and verifying relational models (Hoyle & Panter, 1995). Using SEM, this study holistically examined the causal relationships between search processes (e.g., search tactic selection, system support, satisfaction with search process) and search outputs (e.g., aspectual recall, satisfaction with search results). The model of the SEM analysis is presented in Figure 4-22 (Chapter 4).

Operational definitions and corresponding measures are identified for all the variables in this study. Table 3-5 presents constructs, variables, and operational definitions of this study. Two dimensions of variables were identified in this study, namely "search process" and "search output." Variables are used as evaluation criteria to assess either search process or search output.

In the dimension of search process, user engagement and system support were assessed in terms of search tactic application. In this dimension, four constructs were investigated including: (1) user engagement, (2) system support for search tactics, (3) difficulty in search tactic application, and (4) satisfaction with search process. The construct of user engagement was measured by three evaluation criteria: (1) frequency of search tactics, (2) time spent on applying search tactics, and (3) types of user engagements. These three variables were used to delineate how users engage in search processes by applying different types of search tactics. The construct of system support was evaluated by two evaluation criteria, namely: (1) types of system support and (2) users' perceived system support. Additionally, this study tried to measure difficulty and satisfaction in terms of search tactic application. These seven variables served as evaluation criteria for the search process. In the dimension of search outputs, this study assessed the quality of search results in different aspects. In Task 1, search efficiency was chosen to represent the quality of search outputs. In Task 2, success rate served as a measure of search outputs. In Task 3, two evaluation criteria were adopted including aspectual recall and satisfaction with search results. Operational definition for each output criterion is identified in Table 3-5.

Table 3-5. System support evaluation criteria and associated measures

Dimension	Construct	Criteria	Operational Definition
Search Process	User engagement	Frequency of search tactics	It refers to how many times a specific type of search tactic occurred during a session.
		Time spent on applying search tactics	It refers to how much time a user spent on applying a specific type of search tactic during a session.
		Types of user engagements	It refers to which activities occur in applying a specific type of search tactic during a session.
	System support for tactics	Type of system supports	It refers to in what ways a system supports users' application of a specific type of search tactic during a session.
		Users' perceived system support for search tactics (Task 3)	It refers to level of support a user experiences from a system in applying a specific type of search tactic during a session.
	Difficulty	Difficulty in applying search tactics (Task 3)	It refers to level of difficulty a user experiences while applying a specific type of search tactic during a session.
	Satisfaction with search process	Satisfaction on applying search tactics (Task 3)	It refers to level of satisfaction a user experiences while applying a specific type of search tactic during a session.
Search Outputs	Quality of search outputs	Search efficiency (Task 1)	Time spent on completing a search session (Tullis & Albert 2008)
		Success rate (Task 2)	Percentage of the requested items that users found (Rubin et al. 2008)
		Aspectual recall (Task 3)	It refers to the ratio of aspects of the search topic identified in the documents saved by the subject, to the total number of aspects of the topic (Dumais & Belkin, 2005)
		Satisfaction with search results (Task 3)	It refers to level of satisfaction a user perceives toward search results obtained from an IR process.

3.5. Chapter Summary

In this chapter, the methodology of the study is described. This study recruited 30 undergraduate students and 30 graduate students from a state university in the United States. Subjects were asked to conduct three search tasks – known-item search, specific information search, and exploratory search – using LOC-DL. Data were collected through multiple methods including pre-questionnaires, transaction logs, think-aloud protocols, and post-search questionnaires. Different data analysis techniques were applied to different research questions. For RQ 1, the transaction logs were analyzed to investigate search tactic patterns based on descriptive statistics, kernel regression, transition analysis, and hierarchical clustering. Also, differences in search tactic application were analyzed by task type. For RQ 2, open coding was used to identify types of system supports. Then, user perceptions of system support, difficulty, and satisfaction were measured in applying different types of search tactics. For RQ 3, the relationships between search tactic selections and search outputs were examined based on multiple regression and structural equation modeling.

CHAPTER 4. RESULTS

In Chapter 4, the author investigates three aspects of user-system interactions in digital libraries by answering the proposed research questions. First, users' search tactic application was explored in terms of frequency, time, change in a session, and transition patterns. Second, types of system supports were identified, and perceived system support, difficulty and satisfaction with search tactic application were assessed. Third, the relationships between user search tactic application and search outputs were examined.

4.1. Users' Search Tactic Application

User engagement was investigated by analyzing users' application of search tactics in three different search tasks. Basically, frequency of and time spent on each type of search tactic were explored using descriptive statistics. In addition, this study traced how search tactic selections would change over time within a single session. Moreover, transitions between search tactics were explored to identify users' frequently applied search strategies while using digital libraries.

4.1.1. Frequency of Search Tactics Applied in Three Different Tasks

4.1.1.1. Known-Item Search Task (Task 1 - Frequency)

In the sixty known-item search sessions of this study (Task 1), 652 search tactics were observed in total. On average, each subject applied approximately 10.9 tactics per session. Out of sixty, fifty nine subjects successfully found the requested item within 5 minutes of

the assigned time. As Task 1 was comparatively easier than the other tasks, they were able to find the requested item by applying a smaller number of search tactics. The range of applied tactic frequency was between 4 and 37. In the case of the only failed session, the subject applied 37 tactics. Six subjects completed the task by using only four tactics. They found the requested item by trying a single query attempt. They came up with an appropriate query specific to the task, and the following search results included some relevant items on the first result page. The shortest path to the item from the homepage was: [Creat] \Rightarrow [EvalR] \Rightarrow [AccF] \Rightarrow [EvalI] \Rightarrow "Found". When subjects selected a browsing strategy, it took longer paths than a query creation strategy because they needed to go through several steps of collection categories to reach the relevant item.

Table 4-1 presents the frequency of each type of search tactic applied in Task 1. In this task, AccF was most frequently applied (28.8%). AccF tactics are one of the most essential tactics that enables searchers to move forward different pages or items during the search process. Xplor and EvalR tactics were also frequently employed by showing 21.6% and 16.9% respectively. Approximately 0.92 Creat tactics were observed per session, which indicates less than one new query per session. Eight sessions did not include any Creat tactic, which reveals those sessions relied solely on browsing method in fulfilling the task. Query reformulation tactics were observed 40 times in the sixty sessions (0.67 per session). Out of 52 subjects who created at least one query during a session, only about half of them modified their initial query (48.3%). Twenty eight subjects were able to complete their task with a single query input, without any query modification effort.

Table 4-1. Frequency of each type of search tactic in Task 1

	AccB	AccF	Creat	EvalI	EvalR	Lrn	Mod	Org	Xplor	Total
Total	35	188	55	74	110	4	40	5	141	652
Average ^a	0.58	3.13	0.92	1.23	1.83	0.07	0.67	0.08	2.35	10.87
Percent	5.4%	28.8%	8.4%	11.3%	16.9%	0.6%	6.1%	0.8%	21.6%	100.0%
STD	1.144	1.901	0.420	0.692	1.593	0.249	0.994	0.276	2.174	5.569

^a Average frequency per session

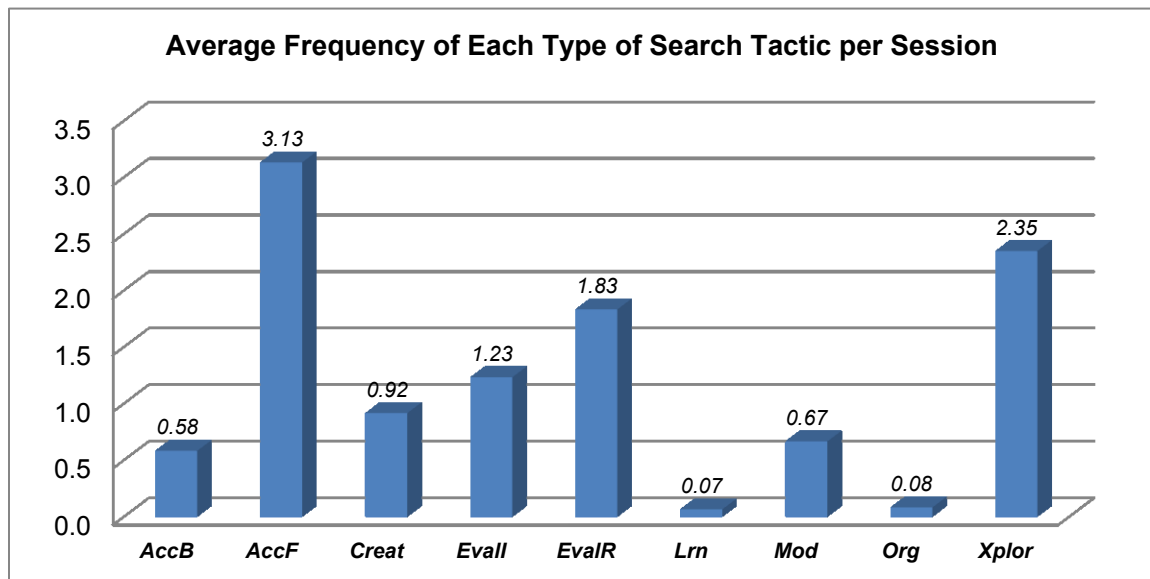


Figure 4-1. Frequency of each type of search tactic in Task 1

Browsing strategy was another frequently applied approach in Task 1. In particular, LOC-DL provides structured browsing categories by topic in individual collections as well as in the homepage. Individual collections also offered categories by topic, time, or other criteria as an access point that assisted users to apply browsing strategies. Eight subjects who relied only on browsing strategies showed iterative patterns of Xplor tactics. For example, S32 completed the task by applying Xplor tactics repeatedly as follows:

[Xplor: Browse by Topic: Select the American Memory collections] \Rightarrow [AccF] \Rightarrow
 [Xplor: Browse by Topic: Select the category of Television/Advertising] \Rightarrow [AccF] \Rightarrow
 [Xplor: Browse by topic: Select the Coca-cola Advertisement Collections] \Rightarrow
 \Rightarrow [AccF] \Rightarrow [EvalI : find the item].

Shifts between browsing and searching were also observed. For example, in five cases, subjects started their session by browsing topic categories, then they switched to creating a query:

[Xplor: Browse by Topic] \Rightarrow [AccF] \Rightarrow [Xplor: Television/Advertising]
 \Rightarrow [Creat: "Coca-cola 1964"] \Rightarrow [EvalR]

Both Lrn (0.6%) and Org (0.8%) tactics were less frequently used. Since Task 1 was relatively easy, subjects rarely encountered any situations in which they needed to apply learning or organizing tactics.

4.1.1.2. Specific Information Search Task (Task 2 - Frequency)

Task 2 was more complicated than Task 1. The sixty subjects applied 1,323 search tactics over the sixty sessions. On average, 22.05 tactics were applied per session. This task asked subjects to find eight specific pieces of information, which are the names of four presidents assassinated and the years of the assassinations. On average, each subject

found approximately 5.1 pieces of information, which is roughly 64% of success rate.

The range of search tactic frequency was between 2 and 37. Interestingly, one of the subjects applied only two tactics in his session, which lasted 164 seconds. He used only one query (Creat) and continued evaluating a long list of search results (EvalR) during the entire session to find the answers. In Task 2, most of relevant information can be obtained directly from surrogates of search results. Thus, subjects were able to achieve the task by evaluating search results (EvalR), even not evaluating individual items (EvalI).

Table 4-2. Frequency of each type of search tactic in Task 2

	AccB	AccF	Creat	EvalI	EvalR	Lrn	Mod	Org	Xplor	Total
Total	169	252	86	159	349	10	158	4	135	1,323
Average ^a	2.82	4.20	1.43	2.65	5.82	0.17	2.63	0.07	2.25	22.05
Percent	12.8%	19.0%	6.5%	12.0%	26.4%	0.8%	11.9%	0.3%	10.2%	100.0%
STD	2.012	2.768	0.667	2.104	2.748	0.637	2.243	0.309	2.803	9.296

^aFrequency per session

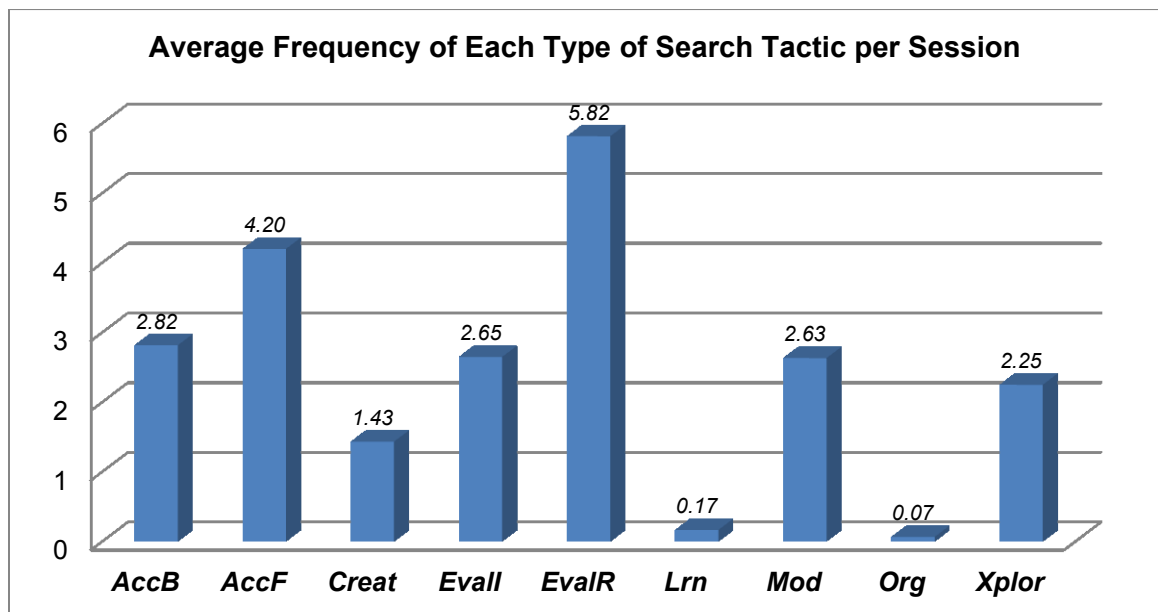


Figure 4-2. Frequency of each type of search tactic in Task 2

Table 4-2 presents the frequency of search tactics observed in Task 2. EvalR tactics were most frequently selected, which makes up 26.4% of all tactics. Unlike the previous task, subjects were able to obtain relevant information from search results because the names of presidents and the assassination years were presented in the search result pages. Therefore, subjects did not necessarily need to access individual items to get answers for the task. AccF and AccB were also frequently applied, 19.0% and 12.7% respectively. These two tactics were used for subjects to move around different pages during the search process. Query formulation and reformulation were applied 6.5% and 11.9%, respectively. In particular, Mod tactics were frequently used when searchers narrowed down search results to get more specific results.

Browsing strategies were less frequently used than search result evaluation strategies in this task. About 10% of the total tactics were Xplor in Task 2. Since Task 2 requested to find specific information, searchers tried to bring up query terms directly indicating the question of the task rather than browsing through topic categories. EvalI comprised 12.0%, which is less than half of EvalR (26.4%). As mentioned above, users could find relevant information from search results, so in many cases, they did not evaluate individual items to get the answers to the search question. Again, minor tactics were least frequently selected: Lrn (0.17%) and Org (0.07%).

4.1.1.3. Exploratory Search Task (Task 3 - Frequency)

Subjects exhibited the most dynamic search tactic patterns in conducting exploratory search tasks. Overall, 3,490 tactics were observed from sixty sessions of Task 3. About 58.17 tactics occurred per each session, and it ranged from 23 to 100. In Task 3, EvalI

tactics were most frequently used by showing 24.4%. To survey various aspects of the topic, users needed to evaluate relevance or usefulness of individual items to complete the search task. Unlike specific information searches, users needed more broad and detailed information on the topic in this task. In many cases, they needed to access individual items and obtained information from the accessed items after judging relevance and usefulness of them. Again, AccF and AccB were frequently used (22.6% and 13.7% respectively) since these two accessing tactics were necessary to move across different pages and items.

Xplor tactics were applied 7.45 times (12.8%) per session, while EvalR tactics were applied 6.38 (11.0%) times. This reveals browsing tactics were a little more frequently selected than result evaluation tactics in Task 3. Similar to the previous tasks, Lrn and Org were the least frequently used tactics.

Table 4-3. Frequency of each type of search tactic in Task 3

	AccB	AccF	Creat	EvalI	EvalR	Lrn	Mod	Obt	Org	Xplor	Total
Total	478	787	80	852	383	8	98	336	21	447	3,490
Average ^a	7.97	13.12	1.33	14.20	6.38	0.13	1.63	5.60	0.35	7.45	58.17
Percent	13.7%	22.6%	2.3%	24.4%	11.0%	0.2%	2.8%	9.6%	0.6%	12.8%	100%
STD	3.75	4.98	0.91	5.92	4.66	0.34	2.01	3.26	0.70	4.66	16.73

^a Frequency per session

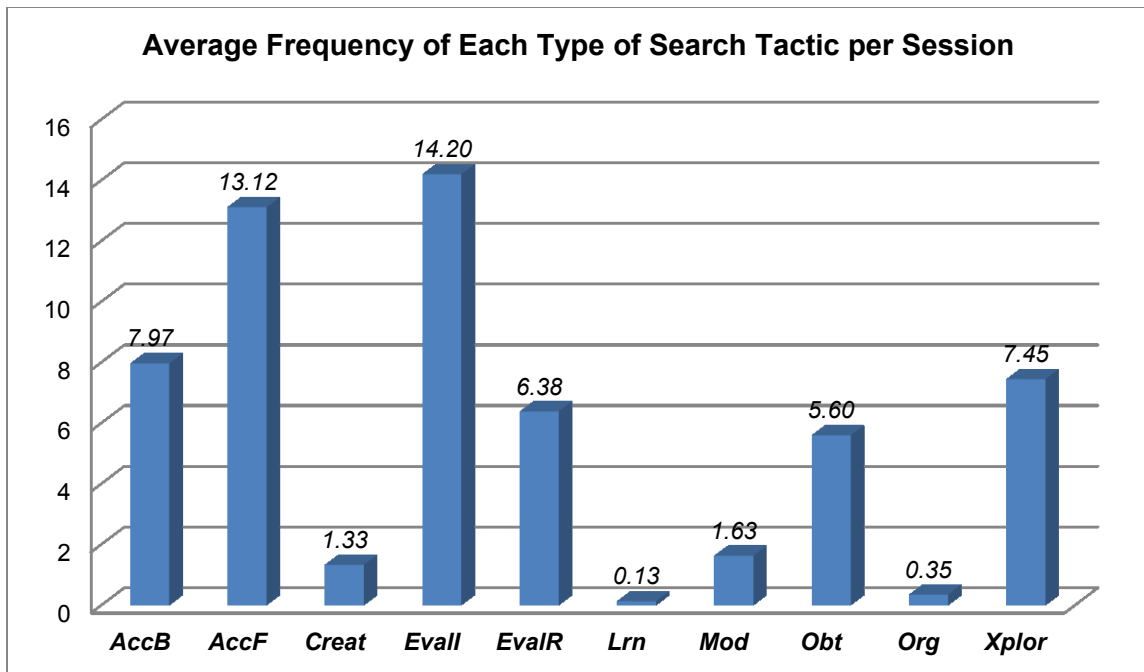


Figure 4-3. Frequency of each type of search tactic in Task 2

4.1.1.4. Search Tactic Frequency Comparison by Task Type

This study further analyzed how search tactic selections differed by task type. Since session lengths varied by session, frequencies of search tactics were standardized by "average number of tactics per minute." Different session lengths were standardized by counting average tactic frequency per minute. Significant differences were found from all seven search tactic types. Minor tactics were excluded in this comparison, including Lrn and Org, because of small numbers of observations. Table 4-4 summarizes the comparison results of search tactic selections by task type.

There were significant differences amongst three task types in applying AccF and AccB. Users applied about 2.1 and 1.8 AccF tactics per minute in Task 1 and Task 3 respectively, while 0.9 AccF tactics in Task 2. This result reveals that they accessed more

frequently to different pages or items in Task 1 and Task 3 than in Task 2. In Task 2, users applied EvalR tactics most frequently and they less frequently accessed (AccF) to individual items as they were able to find relevant information from search results. This explains why there was less frequent application of AccF tactics in Task 2. AccB tactics were most frequently observed in Task 3 because users showed more iterative patterns of search result evaluation or browsing.

As to query creation and reformulation, users applied Creat tactics most frequently (0.731 per minute) in Task 1 compared to the other tasks, $F(1.498, 88.363) = 35.594$; $p < 0.01$. In Task 1, users were likely to use a single query to complete the task in many cases. This explains the higher frequency of Creat tactics in Task 1. Conversely, users more relied on browsing and individual item evaluation tactics in Task 3, whereas relatively less frequent Creat tactics (0.177 per minute). In terms of query reformulation, users most frequently applied Mod tactics in Task 2. It was observed that they were likely to modify previous queries in order to get more specified search results to acquire specific factual information. EvalR tactics were most frequently employed in Task 2 (1.301 per minute). In Task 2, many subjects tried to find relevant information directly from search results rather than individual items, and accordingly EvalR tactics were frequently used throughout the session. On the contrary, EvalI tactics were less frequently used in Task 2. EvalI tactics were most frequently applied in Task 3 by showing about 1.95 per minute. To collect different aspects of information on a particular topic, users need to evaluate a series of related items repeatedly, which results in frequent item evaluation tactics.

Table 4-4. Comparison of frequency of each type of search tactic per minute by task type:
Repeated-measures ANOVA

Tactic	Task ^a	N	Mean	STD	F (Greenhouse-Geisser)	Pairwise comparison ^b
AccF	Task 1	60	2.127	1.222	F(1.498, 88.363) = 35.594 p<0.01	T1-T2** T1-T3 ^{ns} T2-T3**
	Task 2	60	.920	.597		
	Task 3	60	1.817	.743		
AccB	Task 1	60	.268	.436	F(1.885, 111.221) = 56.397 p<0.01	T1-T2** T1-T3** T2-T3**
	Task 2	60	.608	.430		
	Task 3	60	1.087	.506		
Creat	Task 1	60	.731	.532	F(1.229, 75.538) = 50.813 p<0.01	T1-T2** T1-T3** T2-T3**
	Task 2	60	.338	.173		
	Task 3	60	.177	.127		
EvalI	Task 1	60	.891	.483	F(1.792, 105.701) = 95.110 p<0.01	T1-T2** T1-T3** T2-T3**
	Task 2	60	.574	.451		
	Task 3	60	1.952	.804		
EvalR	Task 1	60	1.232	.868	F(1.956, 115.390) = 9.116 p<0.01	T1-T2 ^{ns} T1-T3** T2-T3**
	Task 2	60	1.301	.585		
	Task 3	60	.835	.607		
Mod	Task 1	60	.395	.564	F(1.841, 108.647) = 12.155 p<0.01	T1-T2* T1-T3** T2-T3*
	Task 2	60	.585	.491		
	Task 3	60	.208	.251		
Xplor	Task 1	60	1.503	1.390	F(1.443, 85.142) = 17.500 p<0.01	T1-T2** T1-T3 ^{ns} T2-T3**
	Task 2	60	.498	.594		
	Task 3	60	1.066	.779		

* p<0.05; ** p<0.01; ^{ns} not significant; ^a Task 1: Known-item search; Task 2: Specific information search; Task 3: Exploratory search; ^b Adjustment for multiple comparisons: Bonferroni

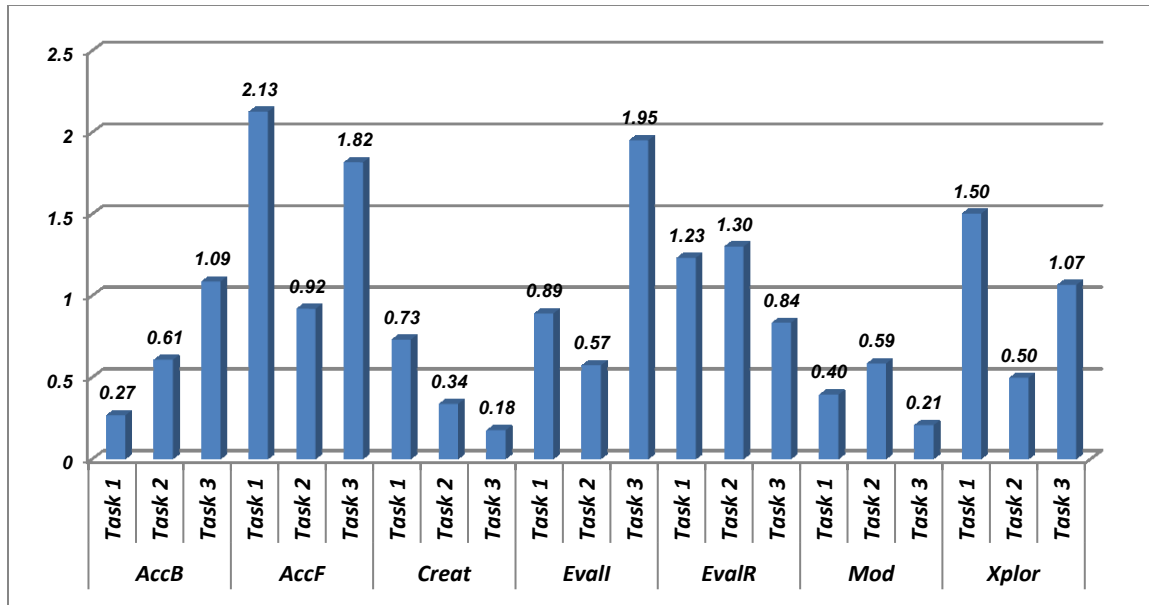


Figure 4-4. Comparison of frequency of each type of search tactic per minute

4.1.2. Time Spent on Applying Search Tactics

This study also investigated users' search tactic application in terms of time spent. To assess users' engagement, how much time users spent on applying each type of search tactic was measured. In this time analysis, AccF and AccB tactics were excluded. Those two accessing tactics indicate changes of pages and items, which were not easily measurable into time using transaction logs.

4.1.2.1. Known-item Search Task (Task 1 - Time Spent)

First, this study investigated searchers' time spent on applying search tactics in Task 1. On average, users completed their Task 1 in 93.9 seconds, and it ranged from 26.77 and 300.00 seconds. The case of 300 seconds is a failed session, which spent up the assigned time, 5 minutes. The longest case of successful sessions was 211.98 seconds. But in most cases (79.7%), users were able to find a relevant item within two minutes. Table 4-5

indicates time spent on applying each type of search tactic in Task 1. Users spent more than half of their Task 1 session for applying EvalR (28.7%) or Xplor (38.4%) tactics. These two tactics represent the major two search strategies, so called "search result evaluation" and "browsing" strategies. Interestingly, users spent a longer time on Xplor than EvalR. Since browsing required several hierarchical steps to reach the final item, it took a longer time than search result evaluation. For query creation and modification, users spent 16.79 seconds (17.9%) and 5.43 seconds (5.8%) respectively. Since users could modify their previous query using search facets, it took less time to reformulate the query than create a new query. Relatively, less time spent on Lrn and Org, 0.52 and 0.19 respectively.

Table 4-5. Time spent on applying search tactics in Task 1

	Creat	EvalI	EvalR	Lrn	Mod	Org	Xplor	Total
Time spent	1007.58	532.24	1618.07	32.65	325.93	11.51	2106.19	5634.17
Average ^a	16.79	8.87	26.97	0.54	5.43	0.19	35.10	93.90
Percent	17.9%	9.4%	28.7%	0.6%	5.8%	0.2%	37.4%	100.0%
STD	11.25	10.83	26.60	2.53	10.43	0.64	34.78	47.82

^a Average time spent per session

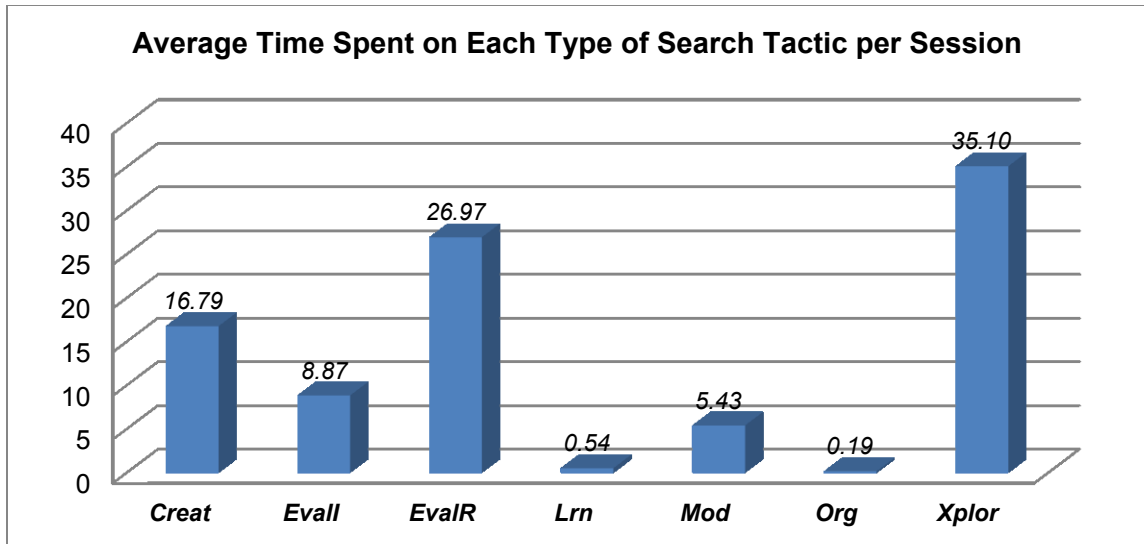


Figure 4-5. Time spent on applying each type of search tactic per session in Task 1

4.1.2.2. Specific Information Search Task (Task 2 - Time Spent)

In Task 2, users completed a session in about 271 seconds on average, which is about 4 minutes 30 seconds. The session lengths ranged between 88 and 300 seconds. Table 4-6 indicates time spent on each type of search tactic in Task 2. Similar to the search tactic frequency results (4.1.1.2), users spent most of their time in evaluating search results, approximately 51 seconds per session (44.6%). Again, in this task, users were able to find relevant information directly from search results. That explains why users spent almost half of the session time in evaluating search results (*EvalR*). The proportions of *Creat* and *Mod* were 8.3% and 13.2% respectively. Relatively less time was spent on *Xplor* tactics (12.1%) compared to *EvalR* tactics (44.6%). Again, less than 1% of session length involved two minor tactics, *Lrn* and *Org*.

Table 4-6. Time spent on applying search tactics in Task 2

	Creat	EvalI	EvalR	Lrn	Mod	Org	Xplor	Total
Time spent	1352.15	3047.67	7276.22	133.76	2146.45	13.74	1978.5	16297.57
Average ^a	22.54	50.79	121.27	2.23	35.77	0.23	32.98	271.63
Percent	8.3%	18.7%	44.6%	0.8%	13.2%	0.1%	12.1%	100.0%
STD	15.55	46.01	53.87	10.31	50.03	1.24	43.75	70.71

^a Average time spent per session

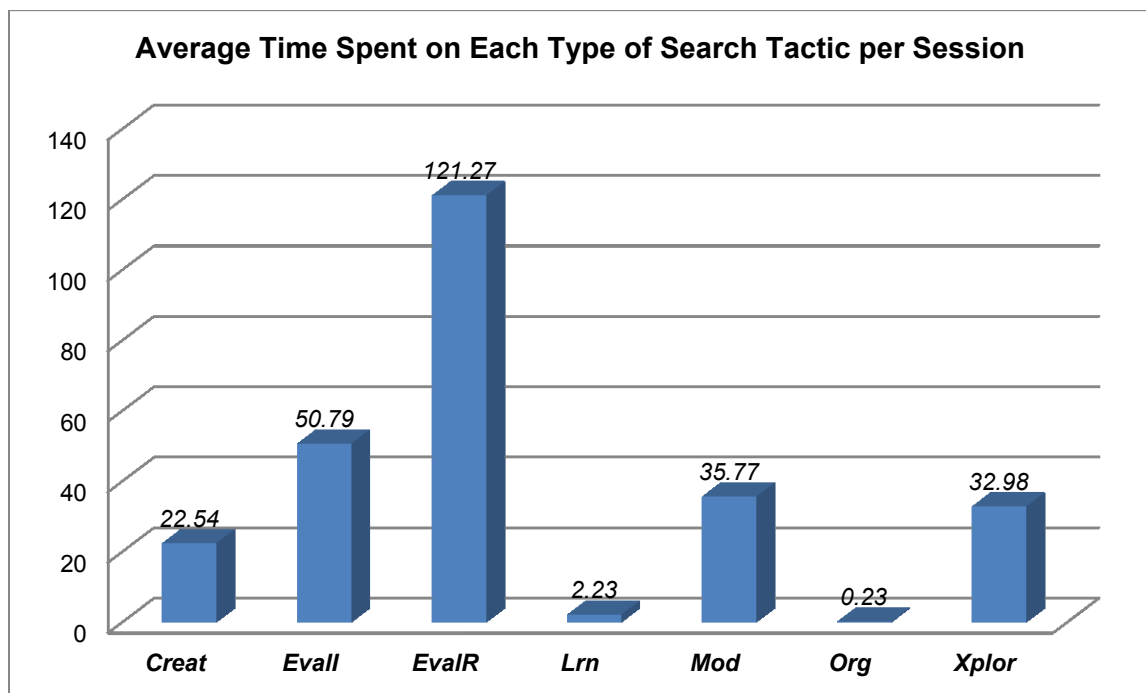


Figure 4-6. Time spent on applying each type of search tactic per session in Task 2

4.1.2.3. Exploratory Search Task (Task 3 - Time Spent)

In Task 3, it took 429.91seconds (about 7 minutes) for subjects to complete a session on average. The range of session lengths was between 263.50 and 480.00 seconds. Table 4-7 shows time spent on applying each type of search tactic in Task 3.

Approximately half of the session time was used to apply EvalI tactics (48.3%). EvalR and Xplor tactics were 18.9% and 15.4% respectively. About 10.3% of the total time was spent on applying Obt tactics. Users accessed individual items mainly through either EvalR or Xplor tactics, and then spent lots of time in evaluating individual items. When finding useful information from EvalI, they saved the information by applying Obt tactics. Since the subjects were asked to gather as many aspects as possible, they went through a series of individual items and reviewed the relevance or usefulness of the accessed items. Relatively less time was associated with query-related tactics, Creat (3.3%) and Mod (2.8%). Because users spent most time in evaluating individual items, time spent on applying query creation and reformulation became a smaller portion of the entire session.

Table 4-7. Time spent on applying search tactics in Task 3

	Creat	EvalI	EvalR	Lrn	Mod	Obt	Org	Xplor	Total
Time spent	862.05	12458.97	4884.61	209.59	719.46	2658.2	37.87	3964.14	25794.89
Average ^a	14.37	207.65	81.41	3.49	11.99	44.30	0.63	66.07	429.91
Percent	3.3%	48.3%	18.9%	0.8%	2.8%	10.3%	0.1%	15.4%	100.0%
STD	10.61	68.93	61.41	11.18	17.79	24.06	1.30	43.09	87.33

^a Average time spent per session

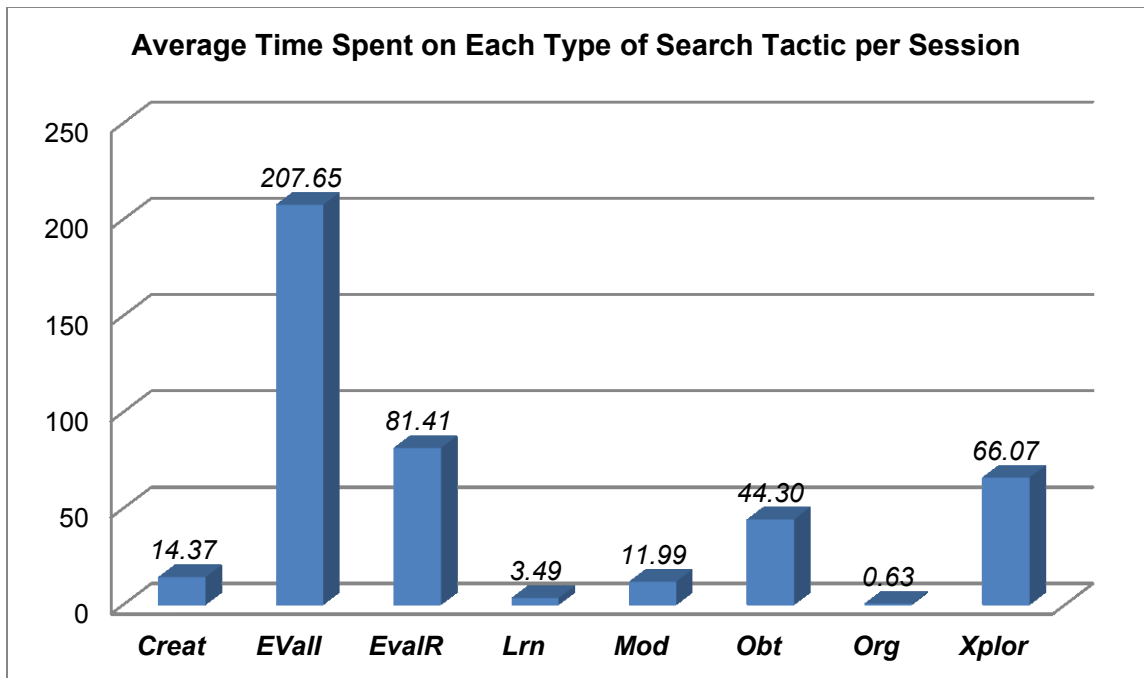


Figure 4-7. Time spent on applying each type of search tactic per session in Task 3

4.1.2.4. Time Comparison by Task Type

This study compared time spent on each type of search tactic by task type. For comparison, we used the percentage data of time spent on search tactics within a session. Significant differences were found from all five types of search tactics. Table 4-8 presents the comparison results of search tactic application by task type.

Table 4-8. Comparison of time spent on search tactics by task type

Tactic	Task ^a	N	Mean	STD	F (Greenhouse-Geisser)	Pairwise comparison ^b
Creat	Task 1	60	.211	.144	F(1.711, 100.942) = 45.165 p<0.01	T1-T2** T1-T3** T2-T3**
	Task 2	60	.100	.126		
	Task 3	60	.032	.023		
EvalI	Task 1	60	.091	.083	F(1.632, 96.260) = 115.787 p<0.01	T1-T2** T1-T3** T2-T3**
	Task 2	60	.186	.168		
	Task 3	60	.469	.138		
EvalR	Task 1	60	.279	.220	F(1.495, 88.217) = 23.333 p<0.01	T1-T2** T1-T3** T2-T3**
	Task 2	60	.491	.369		
	Task 3	60	.174	.124		
Mod	Task 1	60	.044	.070	F(1.444, 85.219) = 25.346 p<0.01	T1-T2** T1-T3 ^{ns} T2-T3**
	Task 2	60	.130	.134		
	Task 3	60	.025	.034		
Xplor	Task 1	60	.346	.310	F(1.505, 88.778) = 21.966 p<0.01	T1-T2** T1-T3** T2-T3 ^{ns}
	Task 2	60	.122	.163		
	Task 3	60	.154	.107		

* p<0.05; ** p<0.01; ^{ns} not significant; ^a Task 1: Known-item search; Task 2: Specific information search; Task 3: Exploratory search; ^b Adjustment for multiple comparisons: Bonferroni

There was a significant difference between task types in applying Creat tactics. The proportion of Creat tactics was high (21.1%) in Task 1 compared to the other tasks, $F(1.711, 100.942) = 45.165$, $p<0.01$. In Task 1, users tended to use a single query to complete the task in many cases. On the other hand, users spent relatively longer time on EvalI tactics in Task 3, so the proportion of Creat turned out to be relatively low (3.2%) in Task 3. In terms of query reformulation, users spent longer time on Mod tactics in Task 2 than the other tasks. This result is consistent with the frequency comparison analysis results. Again, similar to the tactic frequency comparison results, users spent the longest time on EvalR tactics (49.1%) in Task 2, compared to the other two tasks. On the

contrary, the proportion of EvalI tactics was relatively lower in Task 2. In Task 3, users spent the longest time on EvalI application (46.9%).

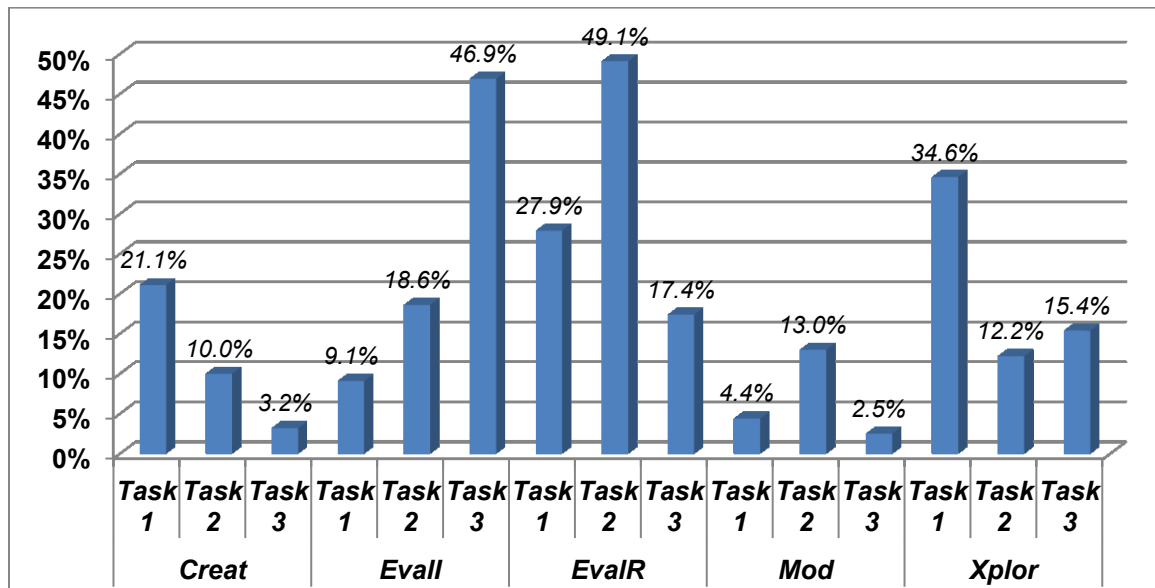


Figure 4-8. Comparison of time spent on search tactics by task type

4.1.3. Search Tactic Probability Change in a Session

This study investigated how search tactic selections change over time within a session.

The author traced the patterns of probabilities of search tactic occurrences in a single session. Using kernel regression, search tactic patterns were estimated for the three types of search tasks.

4.1.3.1. Known-item search task (Task 1)

Figure 4-9 illustrates search tactic probability changes within a single session in Task 1.

Table 4-9 presents the model fits of the kernel regression results measured by R square values and root mean square errors (RMSE). In Task 1, it was observed that users started

search sessions by applying either Creat or Xplor tactic. Creat tactics were selected with higher probability at the starting point than Xplor tactics. Then, the probability of the Creat tactic selection fell down sharply until about the middle of the session. Around the middle of the session, between the period of 0.4 – 0.8 of the session, Xplor and EvalR showed approximately 35% – 45% probabilities respectively. This implies that users applied Xplor or EvalR tactics frequently in the middle phase of a session. Then, users accessed the item after browsing or search result evaluation, and evaluated whether the accessed item was correct as requested from the assigned task. Therefore, the ending phase, approximately from the 0.9 to the end of the session, showed an increased probability of EvalI tactics. As shown in Figure 4-9, the probabilities of search tactic selection differed by phase within a session in Task 1.

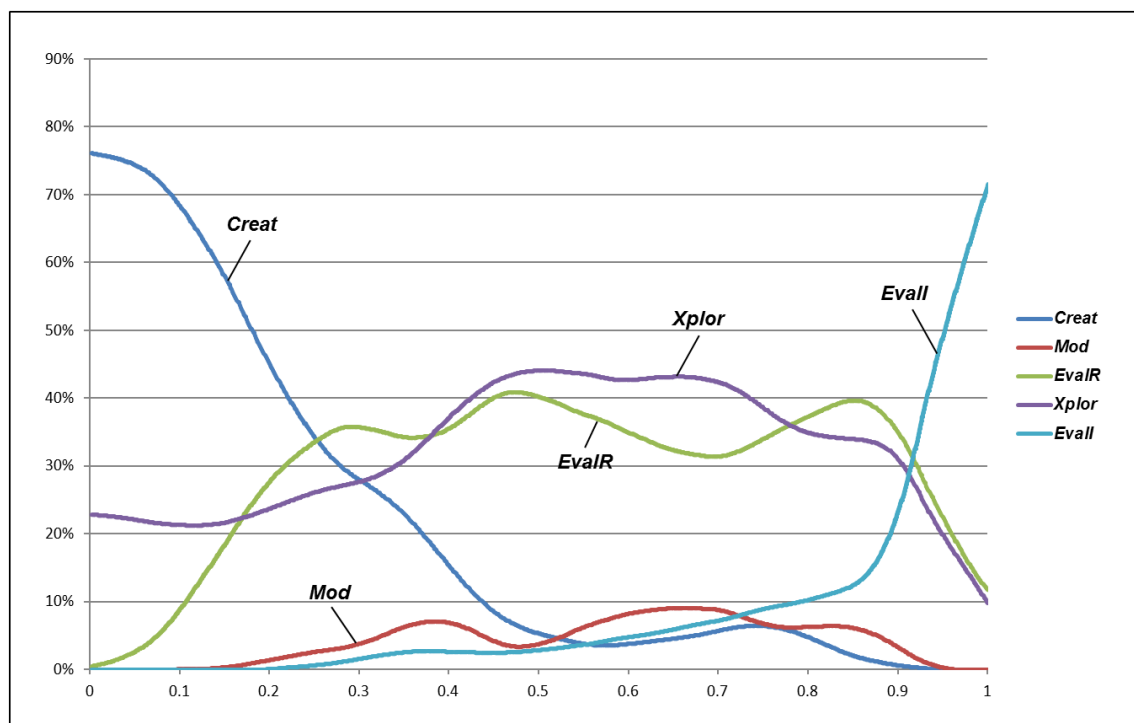


Figure 4-9. Kernel regression of search tactic selection in a session: Task 1

Table 4-9. Model fit of Kernel regression estimation (Task 1)

Tactic	R^2	RMSE
Creat	0.99	0.014
Mod	0.82	0.016
EvalR	0.94	0.030
Xplor	0.93	0.026
EvalI	0.94	0.041

4.1.2.2. Specific information search task (Task 2)

The same analysis was conducted in Task 2. Figure 4-10 illustrates search tactic change patterns in a session in Task 2, and Table 4-10 presents the model fit indices. The starting phase shows that query creation (Creat) and browsing (Xplor) tactics were the two main methods to initiate the session, and query creation was more frequently used than browsing in Task 2. Then, the occurrence of Creat tactics drastically dropped at the region between the starting point and 0.2 of the session. After 0.15 to the end of the session, EvalR tactics frequently occurred showing between 40 and 55 percent. By applying these EvalR tactics, users evaluated many search results related to the topic and acquired relevant information. The probability of EvalI tactics lasted around 20% for the period between 0.2 and 0.7 of the session while there was a temporal hump up to 30% around the 0.9 point area of the session. Also, the occurrence of Mod tactics were observed steadily across the session, which was between 10% and 15%.

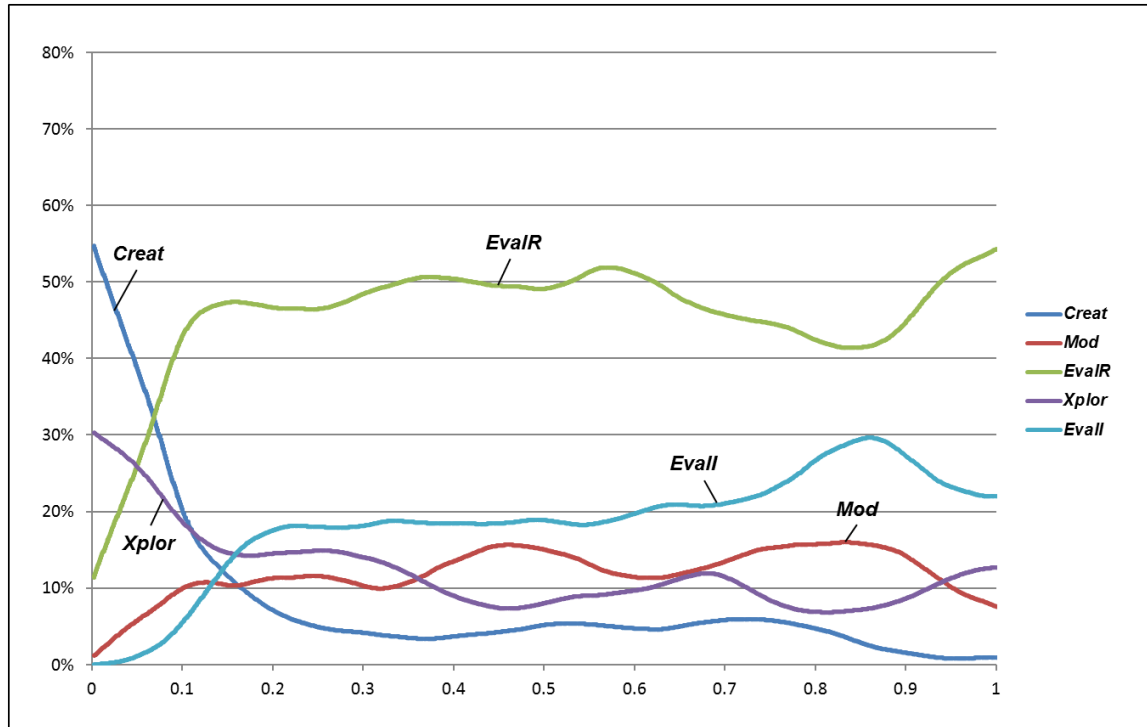


Figure 4-10. Kernel regression of search tactic selection in a session: Task 2

Table 4-10. Model fit of Kernel regression estimation (Task 2)

Tactic	R^2	RMSE
Creat	0.91	0.037
Mod	0.62	0.028
EvalR	0.78	0.048
Xplor	0.90	0.018
EvalI	0.89	0.025

4.1.2.3. Exploratory search task (Task 3)

Figure 4-11 illustrates the changes of search tactic selection probabilities within a session in Task 3. Table 4-11 reports on the model fits of the regression models. An R^2 value was relatively low for EvalR, which is 0.29. Other than that, the kernel regression achieved adequate model fits by showing R^2 values over 0.6.

Similar to the previous tasks, users initiated their session by employing either Xplor or Creat tactics. In this task, Xplor tactics were more frequently selected than Creat tactics in the beginning of the session. Both Xplor and Creat drastically plunged down right after the starting point. In particular, Creat tactics were less frequently selected after the 0.15 point of the session. The probability of Xplor tactic occurrence was roughly between 10% and 15% after 0.15 of the session.

EvalI showed a high probability after 0.15 of the session. From 0.15 to the end of the session, the probability of EvalI lasted approximately fifty percent. The probability of Obt tactics reached the highest point at approximately 0.17 of the session, and then lasted around 10% with a slight decrease throughout the session. Overall, users showed a high EvalI tactic probability throughout the session.

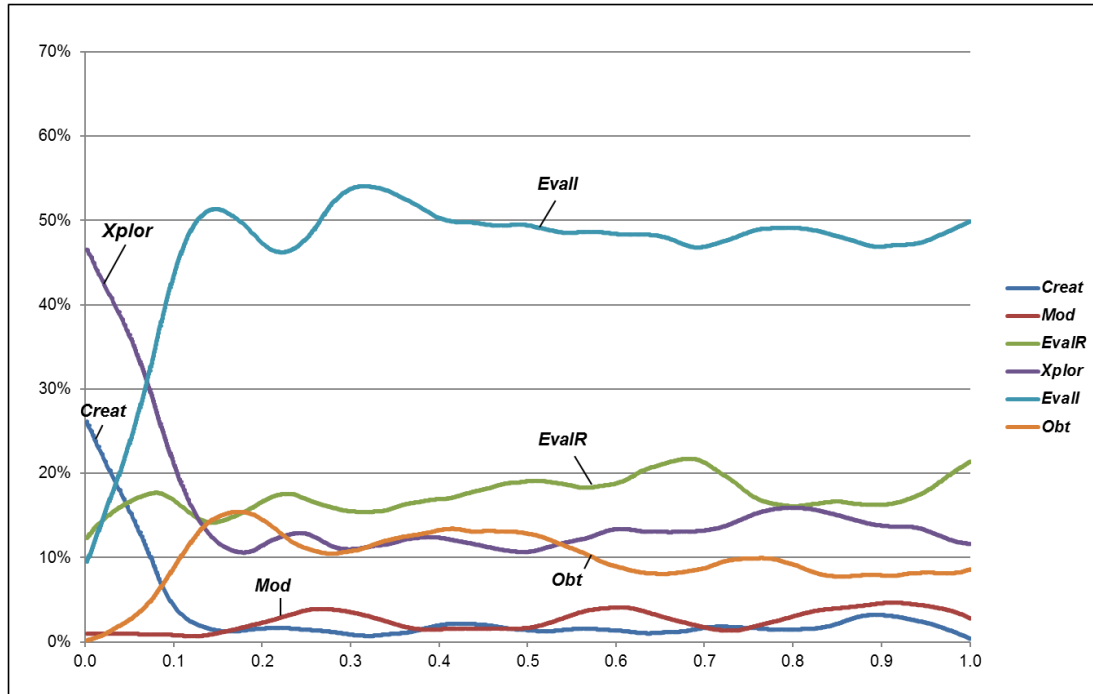


Figure 4-11. Kernel regression of search tactic selection in a session: Task 3

Table 4-11. Model fit of Kernel regression estimation (Task 3)

Tactic	R ²	RMSE
Creat	0.81	0.028
Mod	0.67	0.016
EvalR	0.29	0.041
Xplor	0.71	0.043
EvalI	0.88	0.056
Obt	0.64	0.030

4.1.4. Transitions in Search Tactics

In order to explore the patterns of user engagement, this study mapped out transitions between search tactics. Based on transition analysis, common patterns of search tactic selection were identified, which delineates users' search strategies applied in different search task situations. The probabilities of tactic transitions were calculated to come up with search tactic transition models based on Markov switching analysis.

4.1.4.1. Known-item Search Task (Task 1)

To investigate transitions between tactics, a directed matrix of search-tactic transitions was created. Transitions from one tactic to another were tabulated in the transition matrix presented in Table 4-12. Obviously, most frequently observed transitions occurred between tactics with high frequency. There were relatively many transitions between accessing tactics (AccF and AccB) and result evaluation and browsing tactics (EvalR and Xplor). The most two frequent transitions were between Xplor and AccF: [Xplor] \Rightarrow [AccF] (N=116) and [AccF] \Rightarrow [Xplor] (N=105). These transitions between Xplor and AccF represent repeated uses of browsing tactics during the search process. It was

observed that users needed to browse three or four times consequently to access different depths of topic categories or item lists to find a relevant item when they using a browsing method. Accordingly, frequent transitions between Xplor and AccF tactics were observed to walk through different hierarchical paths to reach the relevant item. Creat and Mod tactics led to EvalR since the next tactic after a query input was to evaluate search results.

Table 4-12. Frequency and probability matrix of search-tactic transitions in Task 1^a

\To From	AccB	AccF	Creat	EvalI	EvalR	Lrn	Mod	Org	Xplor
AccB	1 (0.03)			1 (0.03)	13 (0.37)		3 (0.09)		17 (0.49)
AccF			5 (0.03)	73 (0.39)	1 (0.01)	4 (0.02)			105 (0.56)
Creat					55 (1.00)				
EvalI	11 (0.61)	7 (0.39)							
EvalR	5 (0.05)	65 (0.61)					36 (0.34)	1 (0.01)	
Lrn	4 (1.00)								
Mod					40 (1.00)				
Org					1 (0.20)				4 (0.80)
Xplor	14 (0.10)	116 (0.83)	5 (0.04)				1 (0.01)	4 (0.03)	

^a Values in parentheses indicate probabilities of transitions between tactics

Figure 4-12 highlighted two main paths to reach a relevant item. At the beginning point, users selected either query creation or browsing tactic. Out of sixty sessions, forty-four users started the search with query creation (Creat: 73.3%) while sixteen did with browsing (Xplor: 26.7%). As shown in Figure 4-12, search result evaluation (EvalR) and browsing (Xplor) were the two major search tactics in constructing users' search tactic transition patterns in Task 1.

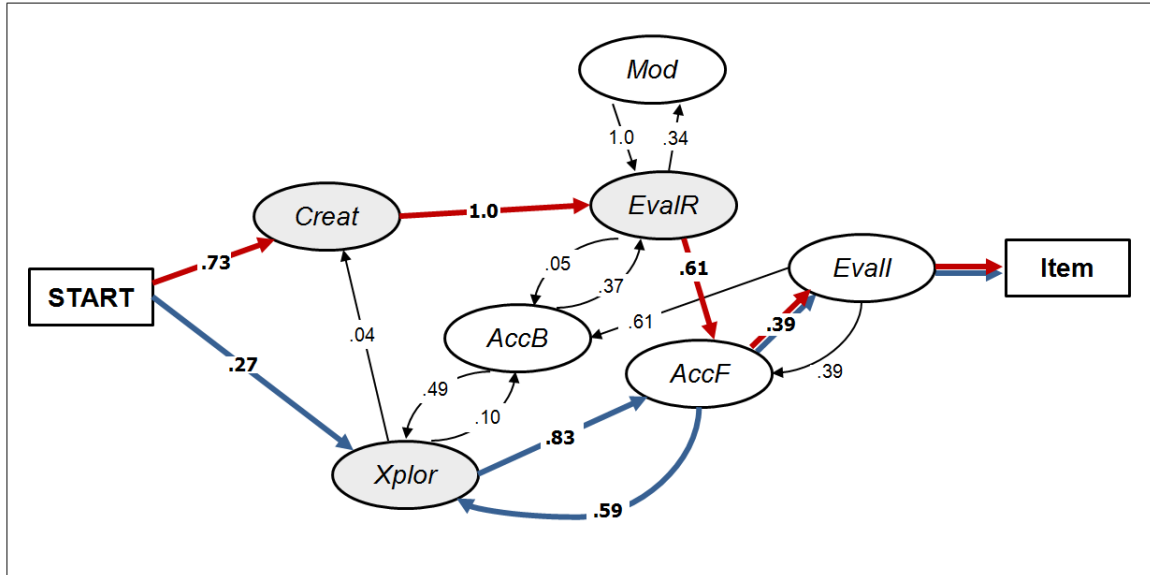


Figure 4-12. Search tactic transition model in Task 1 (minor tactics or transitions less than 5 were excluded in the diagram)

The red line path represents a search result evaluation strategy. The search result evaluation strategy can be illustrated into two ways. The following (a-1) pattern shows the simplest path to an item with a single query attempt, while (a-2) represents a query result evaluation strategy including query reformulation(s).

(a-1) [Start] \Rightarrow [Creat] \Rightarrow [EvalR] \Rightarrow [AccF] \Rightarrow [EvalI] \Rightarrow [Task completion]

(a-2) [Start] \Rightarrow [Creat] \Rightarrow [EvalR] \Rightarrow [Mod (can be multiple)] \Rightarrow [EvalR] \Rightarrow [AccF] \Rightarrow [EvalI] \Rightarrow [Task completion]

The blue line path delineates a browsing strategy. In this path, users usually started a session by browsing topic categories in the homepage, and continued browsing subordinate categories or a list of items until accessing to a relevant item. The path (b)

shows a typical tactic transitions of repeated browsing and following item evaluation to complete the task:

(b) [Start] \Rightarrow [Xplor] \Rightarrow [AccF] \Rightarrow ... several rounds of browsing and accessing...
 \Rightarrow [AccF] \Rightarrow [EvalI] \Rightarrow [Task completion]

Of course, there were shifts between search strategies. For example, in five search sessions, we observed shifts of search strategies from browsing to query formulation ([Xplor] \Rightarrow [Creat]).

4.1.4.2. Specific Information Search Task - Task 2

A directed transition matrix was generated for Task 2 (Table 4-13). As EvalR was most frequently observed in this task, there were numerous transitions between "EvalR and AccF," "EvalR and AccB," "EvalR and Creat," and "EvalR and Mod." That is, EvalR played a key role in the search process of Task 2. In particular, EvalR tactics led to many query modification (136 transitions). This type of transition reveals that users tried to change search terms frequently when search results were not satisfactory.

Transitions among EvalR, AccF, and EvalI were also frequently observed. Another route to acquire relevant information was from evaluating individual items, so patterns of [EvalR] \Rightarrow [AccF] \Rightarrow [EvalI] were also applied frequently. The proportion of browsing strategies was less recurrent compared to query-based strategies. Xplor tactics were linked to AccF and then EvalI. Also, about 19% of Xplor tactics were switched to Creat tactics.

Table 4-13. Frequency and probability matrix of search-tactic transitions in Task 2^a

\To From	AccB	AccF	Creat	EvalI	EvalR	Lrn	Mod	Org	Xplor
AccB	2 (0.01)		3 (0.02)	3 (0.02)	102 (0.61)	1 (0.01)	14 (0.08)		43 (0.26)
AccF			11 (0.04)	156 (0.62)	7 (0.03)	6 (0.02)	1 (0.00)		71 (0.28)
Creat					85 (1.00)				
EvalI	114 (0.78)	26 (0.18)	3 (0.02)				3 (0.02)		
EvalR	28 (0.09)	135 (0.43)	6 (0.02)			2 (0.01)	136 (0.44)	4 (0.01)	
Lrn	6 (0.67)	3 (0.33)							
Mod	2 (0.01)	3 (0.02)			148 (0.96)	1 (0.01)			
Org					4 (1.00)				
Xplor	16 (0.13)	85 (0.66)	24 (0.19)				3 (0.02)		

^a Values in parentheses indicate probabilities of transitions between tactics

Figure 4-13 illustrates paths to reach relevant items that are identified from the transition analysis. Similar to Task 1, users employed either query creation or browsing as a session starting method. Out of sixty sessions, thirty nine users started with query creation (Creat: 65.0%) while twenty one started with browsing (Xplor: 35.0%). Search result evaluating strategy was most frequently used. In particular, iterative loops of search result evaluation were often used in Task 2.

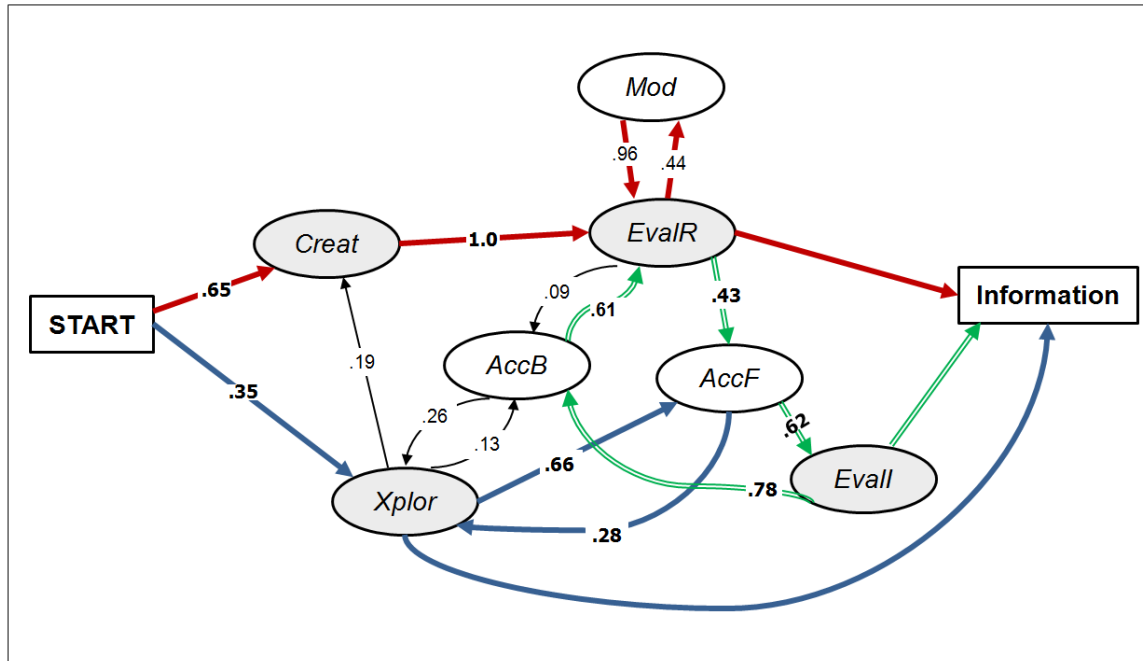


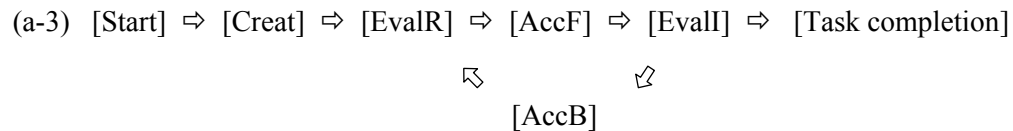
Figure 4-13. Search tactic transition model in Task 2 (minor tactics or transitions less than 10 were excluded from the diagram)

Common patterns of sequential tactics were identified from the transition analysis. The red line path represents a search result evaluation strategy. The search result evaluation strategies are illustrated into two paths, which are the shortest path to an item with a single query attempt (a-1) and a path with query reformulation(s). In cases of (a-1), users were able to find relevant information from the initial search results without any query reformulation effort, which is therefore more efficient.

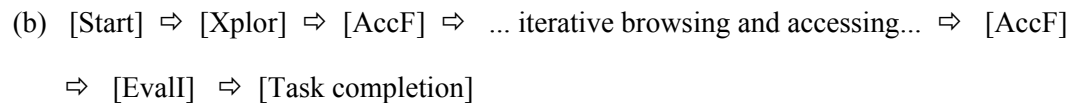
(a-1) [Start] \Rightarrow [Creat] \Rightarrow [EvalR] \Rightarrow [Task completion]

(a-2) [Start] \Rightarrow [Creat] \Rightarrow [EvalR] \Rightarrow [Mod (can be multiple)] \Rightarrow [EvalR] \Rightarrow [AccF]
 \Rightarrow [EvalI] \Rightarrow [Task completion]

In addition, iterative loops were also observed between EvalR and EvalI tactics. The green path of Figure 4-13 represents iterative loops between EvalR and EvalI. In this loop, users repeatedly evaluated search results, and accessed and evaluated individual items to find relevant information. Accordingly, iterative paths occurred between EvalR and EvalI. AccF and AccB tactics were used to connect EvalR and EvalI in those loops as shown in the path of (a-3) below.



The blue line path delineates browsing strategies, and these browsing paths showed iterative patterns. The iterative browsing strategies have a typical pattern of:



However, these browsing strategies were infrequently used in Task 2.

4.1.4.3. Exploratory Search Task - Task 3

Task 3 exhibits more complex transition patterns compared to the other task types. Table 4-14 presents frequencies and probabilities of transitions in search tactics observed from sixty Task 3 sessions. The most frequent transition was from AccF to EvalI (N=562). The transition from AccF to EvalI represents that a user accessed an item and then evaluated its relevance and usefulness. The second most frequent transition was from

Xplor to AccF (N=347), which indicates that a user browsed topic categories and then accessed a specific item or collection. Transitions of "EvalI \Rightarrow Obt", "EvalI \Rightarrow AccB, " and "EvalR \Rightarrow AccF" also were frequently observed and ranked third, fourth, and fifth, respectively.

Table 4-14. Frequency and probability matrix of search-tactic transitions in Task 3^a

\To From	AccB	AccF	Creat	EvalI	EvalR	Lrn	Mod	Obt	Org	Xplor
Accb	20 (0.04)		6 (0.01)	77 (0.16)	182 (0.38)		19 (0.04)			175 (0.37)
Accf			12 (0.02)	562 (0.72)	1 (0.00)	1 (0.00)	2 (0.00)	4 (0.01)	1 (0.00)	203 (0.26)
Creat					77 (0.963)					3 (0.038)
EvalI	293 (0.36)	191 (0.23)	4 (0.01)			4 (0.01)	2 (0.00)	320 (0.39)	1 (0.00)	8 (0.01)
EvalR	47 (0.13)	224 (0.61)				3 (0.01)	67 (0.18)	9 (0.03)	16 (0.04)	
Lrn	3 (0.38)			3 (0.38)			1 (0.13)	1 (0.13)		
Mod					98 (1.000)					
Obt	67 (0.21)	24 (0.07)	1 (0.00)	208 (0.64)	6 (0.02)					19 (0.06)
Org				2 (0.10)	16 (0.76)					3 (0.14)
Xplor	49 (0.11)	347 (0.79)	33 (0.08)				7 (0.02)	1 (0.00)	3 (0.01)	

^a Values in parentheses indicate probabilities of transitions between tactics

More importantly, iterative patterns were frequently observed in both browsing and search result evaluation strategies. Higher order Markov chains were applied to trace the unique iterative patterns of search tactic transitions. In this task, a probability of each shift was calculated in consideration of all the past tactics occurred before a specific tactic. Two most frequently observed search strategies were identified and probability was calculated for each transition (Figure 4-14).

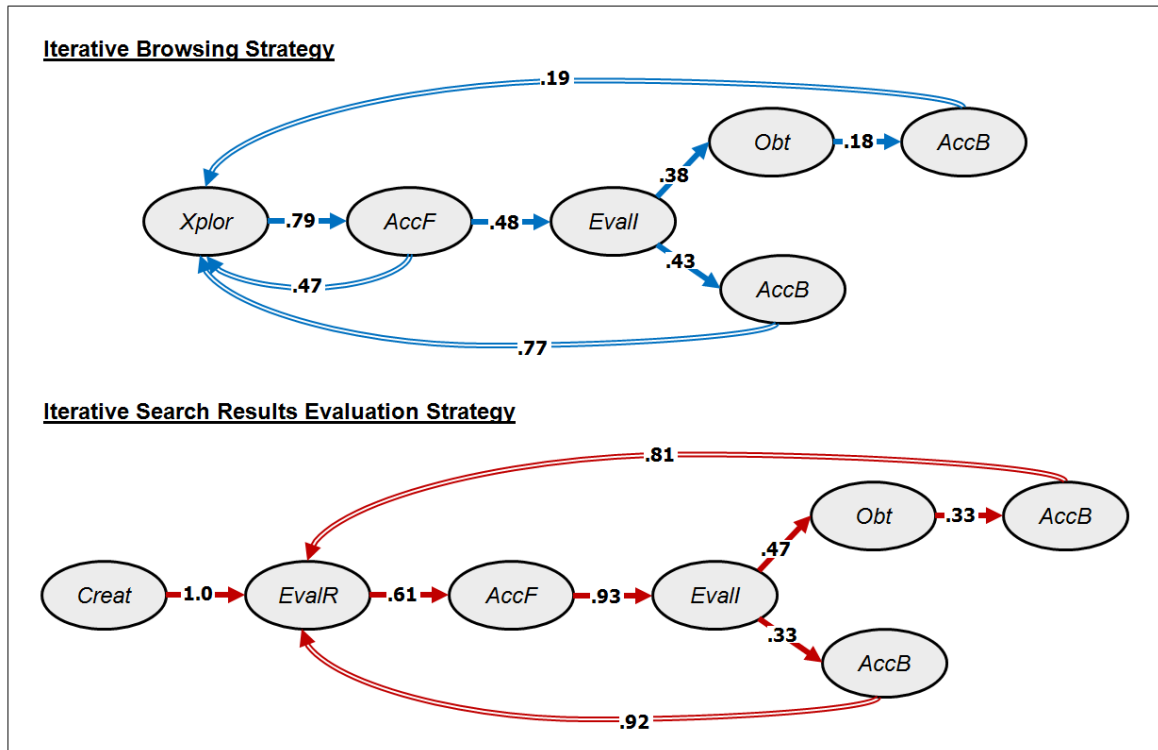


Figure 4-14. Two most frequently applied search strategies identified from the transition analysis: Iterative browsing and iterative search results evaluation

First, the iterative browsing strategy can be represented by a sequential pattern in which a user browses and evaluates a series of items by exploring different topic categories until she/he is satisfied or quits. From the observations of the Task 3 sessions, we found that users accessed items from browsing categories and evaluated whether to obtain the accessed items based on relevance judgment. Once they decided relevant to the topic, they typically obtained the entire item or pieces of information from the item. Then, they either went back to the previous browsing categories for further exploration or quit the session. Otherwise, they accessed back to the previous browsing categories without obtaining any information for further exploration of categories or a list of items.

Second, the iterative result evaluation strategy showed a sequential tactic pattern where a user evaluates search results iteratively to access and to find relevant information. Users tended to evaluate search results from the top and accessed an item based on the evaluation of each search result. In this strategy, users usually accessed back to the list of search results and evaluated the rest of the search results several times until they found sufficient relevant information. In this pattern, EvalR tactics occurred repeatedly, and users determined whether any item is worth selecting for detailed evaluation (EvalI). The two sequences presented in Figure 4-14 show the typical sequences of tactics for the two most frequently applied search tactics – iterative browsing strategy and iterative search result evaluation strategy.

4.1.5. Clustering Search Sessions by Search Tactic Selection

This study further analyzed the characteristics of search sessions based on users' search tactic selection. Search sessions were clustered to identify sessions with similar search tactic patterns.

To classify search sessions based on users' selection of search tactics, hierarchical clustering and multi-dimensional scaling were employed sequentially. The author conducted a hierarchical clustering analysis based on Minkowski distance and Ward's method. Dendrograms were drawn to determine the clusters of sessions (Figure 4-15, 4-17, and 4-19), and three or four groups were identified for each type of search task.

Firstly, three groups of search sessions were identified in Task 1 based on hierarchical clustering. A two dimensional MDS map was created based on Euclidian distance (Figure 4-16). A stress value of 0.08 and an R^2 of 0.97 were obtained. The identified three groups

are: query and result evaluation oriented sessions (Group 1); browsing oriented sessions (Group 2); and combination of browsing and result evaluation sessions (Group 3). The users of Group 1 relied on query creation and following search result evaluation while few browsing tactics were observed during the sessions. The Group 2 users applied Xplor tactics more frequently. The Group 3 sessions, which is located in the middle of the MDS map, included both query-related tactics and browsing tactics.

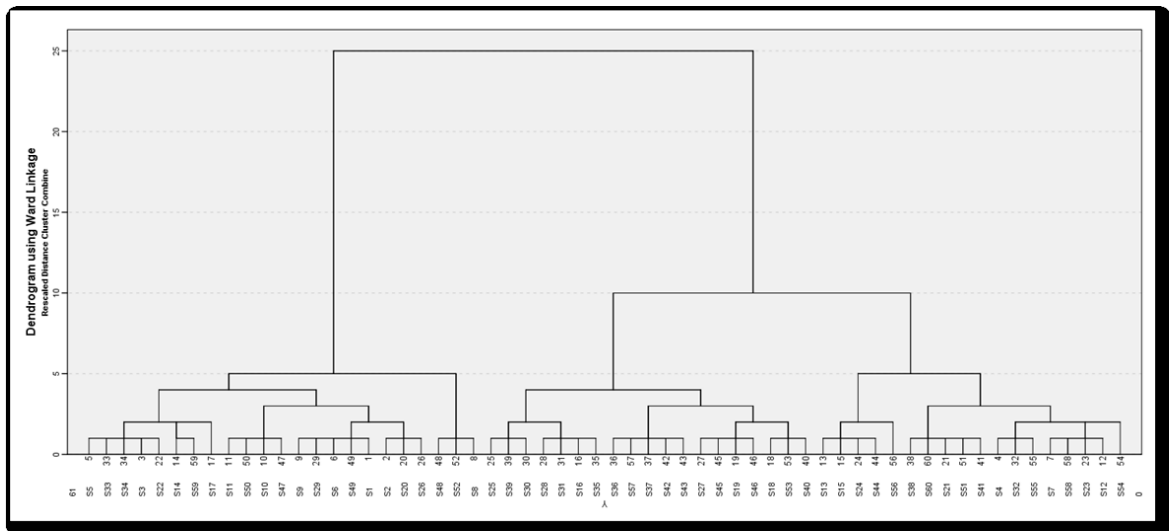


Figure 4-15. Clustering analysis of search sessions in Task 1

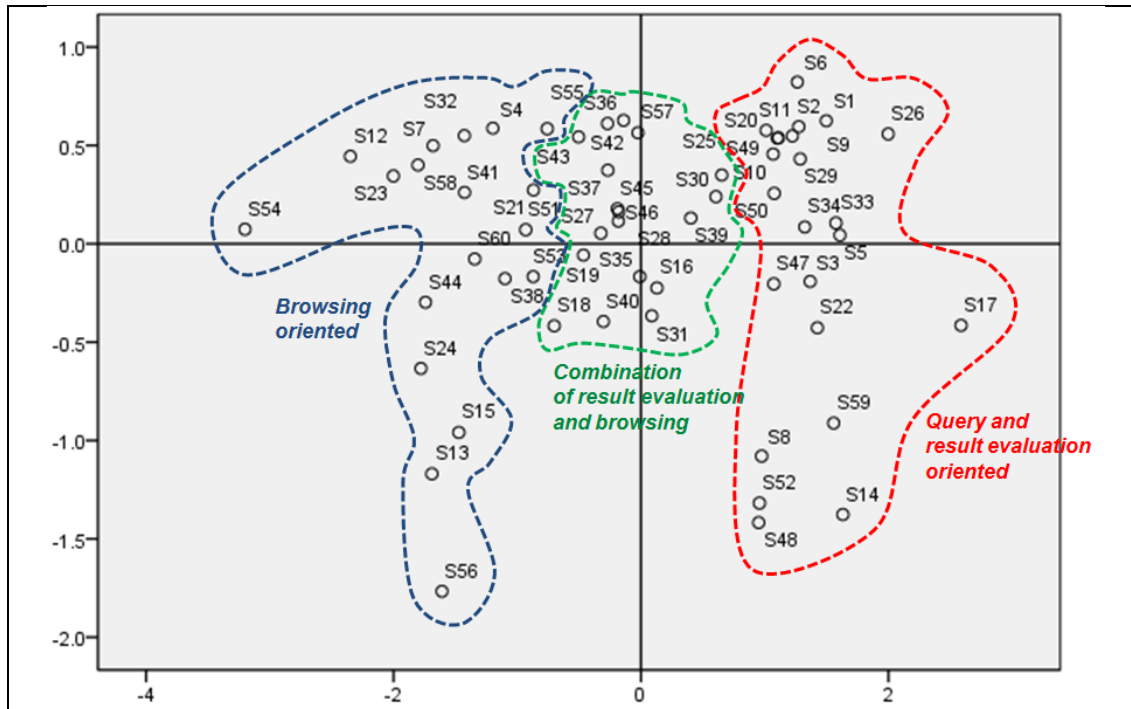


Figure 4-16. MDS map of search sessions in Task 1

Secondly, Task 2 sessions were also clustered based on search tactic selection. Four groups were identified from hierarchical clustering. All the sixty Task 2 sessions were projected on the two-dimensional space based on Euclidean distance (Figure 4-18). A stress value of 0.10 and an R^2 of 0.95 were achieved. The identified four groups are: query creation oriented sessions (Group 1); iterative research evaluation sessions (Group 2); iterative result evaluation with frequent query modification sessions (Group 3); and browsing oriented sessions (Group 4). In Group 1, users were likely to complete their search session with query creation and result evaluation, while less iteration of result evaluation. Group 2 involved typical patterns of iterative search result evaluation strategies, with repeated uses of AccF, AccB, and EvalR. Group 3 was a variant of Group 2 with frequent uses of Mod tactics. Users in Group 4 showed frequent Xplor tactic application compared to other groups.

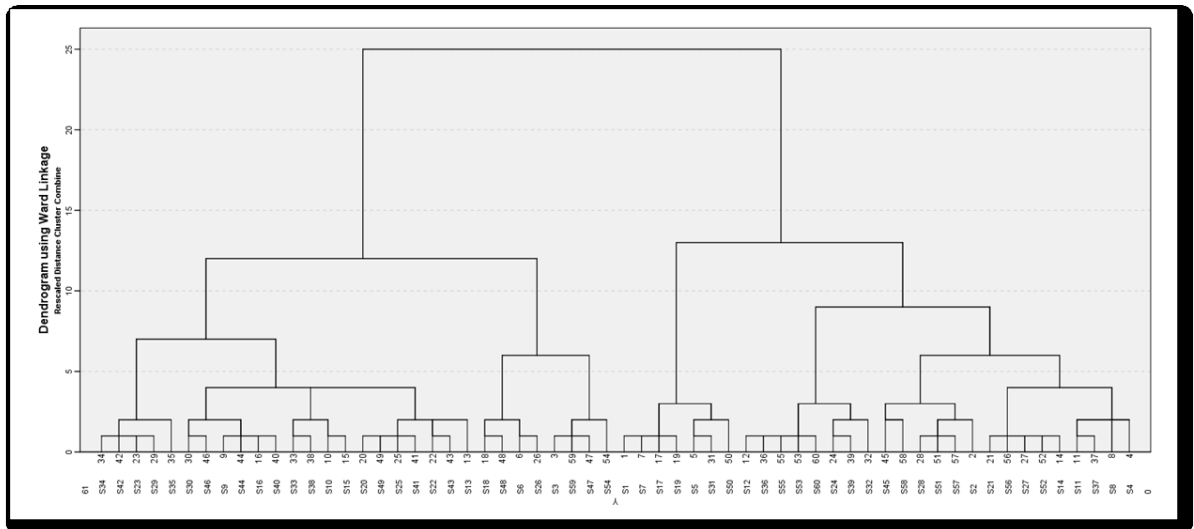


Figure 4-17. Clustering analysis of sessions in Task 2

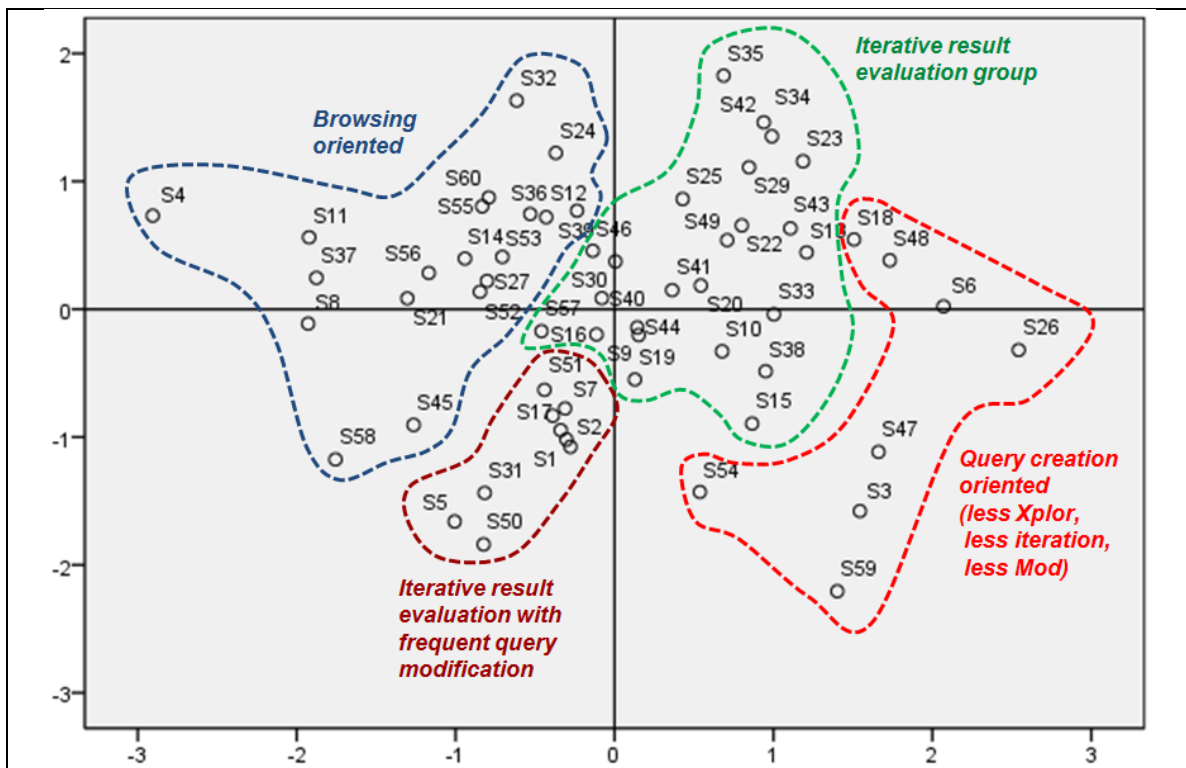


Figure 4-18. MDS map of search sessions in Task 2

Thirdly, the sixty sessions of Task 3 were classified into three groups. Multidimensional scaling exhibited high stress in the two-dimensional space, which is 0.13 ($R^2=0.92$). The groups identified are: iterative browsing oriented sessions (Group 1); iterative result

evaluation oriented sessions (Group 2); and combination of browsing and result evaluation sessions (Group 3). In this way, search sessions had different characteristics by search tactic selection patterns, and sessions were clustered by their similarities in search tactic selection.

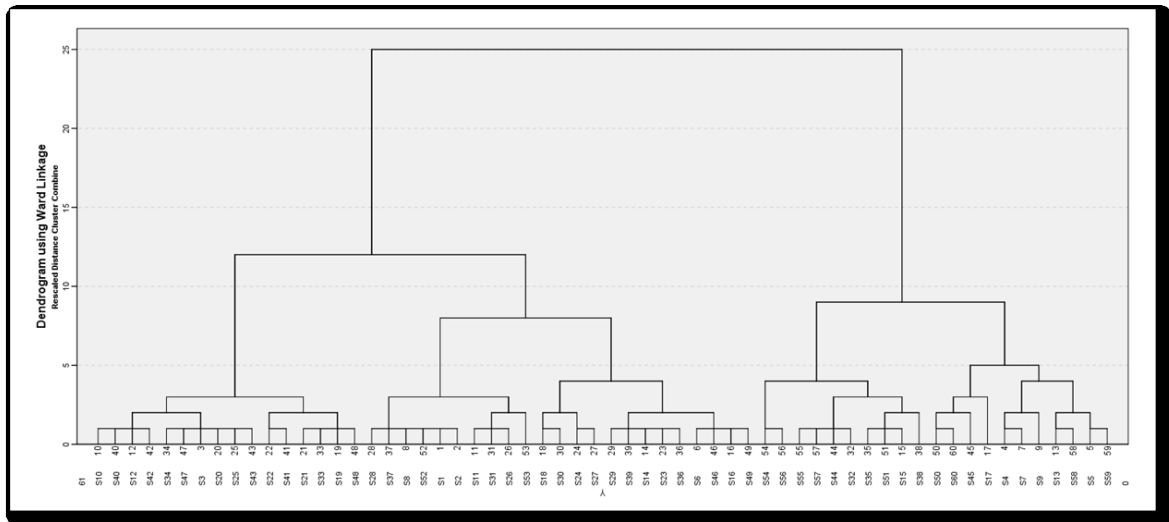


Figure 4-19. Clustering analysis of sessions in Task 3

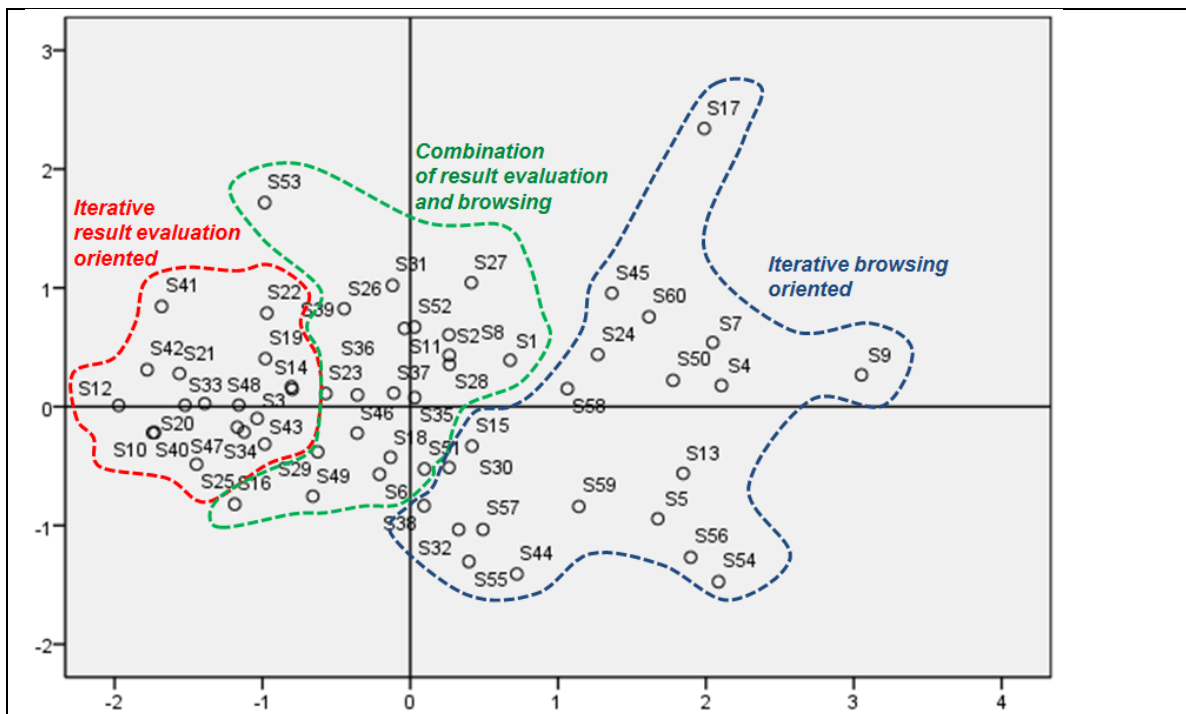


Figure 4-20. MDS map of sessions in Task 3

4.2. System Support for User Search Tactics

This section (1) investigates types of system support for users' search tactic application in LOC-DL and (2) measures perceived system support for each type of search tactic. In addition, this study investigated users' perceived difficulty in and satisfaction with applying different types of search tactics.

4.2.1. Types of System Support for Search Tactics

As the first step to investigate system support, this study analyzed a variety of system features provided in LOC-DL. This study identified different types of system features for each type of search tactic based on the observation of search processes. Minor tactics were excluded in this analysis due to their infrequent occurrences, such as Org and Lrn.

4.2.1.1. Types of System Support for Creat tactics

LOC-DL has limited system features for query creation compared to commercial search engines or online databases. The advanced search function had limited availability in subordinate collections. Some collections provided an advanced search function while some did not. The basic search box was most frequently used for applying Creat tactics in LOC-DL because it was the only option to submit a query in many collections as well as the homepage. Table 4-15 summarizes system features for Creat tactics provided by LOC-DL.

Table 4-15. Types of system support for Creat tactics in LOC-DL

System support	Related system features	Example
Support users to send a query easily and efficiently	Basic search	S9: She typed a query, "president assassination", and sent it to the system using the basic search box in the homepage.
Support users to manage their search queries effectively	Advanced search	S29: He used an advanced search function to limit search results in a specific time period.
	Boolean search	S38: She input a query "coca-cola AND 1964 and ad"
Support users to select more relevant terms	Query suggestion/ Query expansion	S3: She selected a query "coca cola television" from the suggestions by the search box.
Support users to search from multiple collections at one time	Collection selecting options/ Federated search	S7: He checked two checkboxes of related collections in his search of multiple collections.
Give users an option to select an appropriate query matching method	Retrieval matching option	S10: She set to "match any of these words" in the search box.
Support users to input correct format of terms	Spelling correction	S38: She found a correct spelling of "assassination" instead of "assasination".

Basically, LOC-DL offered a basic search box in the homepage, which enabled users to search from all subordinate collections of LOC-DL. Both Boolean operators and query suggestion were provided in the basic search box. Advanced search functions were selectively available in individual collections, such as "Historic Newspapers" and "American Memory Collections." Query suggestion was frequently used as it was provided in the main search box of the homepage. In addition, LOC-DL provided an option to selectively search subordinate collections. LOC-DL had a separate search help

guide, named as "Search Help". It provided detailed search tips with examples, but it was infrequently used. Only twenty one visits of Help pages were observed from the entire search logs of 180 sessions in this study.

4.2.1.2. Types of System Support for EvalR Tactics

Once users sent a query to the system, they engaged in search results. LOC-DL offered several different system features for EvalR tactics. Table 4-16 summarizes types of system support and associated system features. Like other search engines, LOC-DL presented the number of retrieved search results by default. LOC-DL provided descriptive information about the retrieved items, such as a title, short description, source, format, and thumbnail. Elements of meta-information differed by collections. For example, American Memory, which contains historical collections in LOC-DL, presented only titles and collection names, while Historic Newspapers collections showed titles, thumbnails, and dates. These surrogates or meta-information elements supported users to examine different aspects of search results, which are necessary to judge the relevance of an item. In addition, highlighting search terms helped users more efficiently evaluate the relevance of collections or items.

Table 4-16. Types of system support for EvalR tactics in LOC-DL

System support	Related system features	Example
Support users to identify the amount of retrieved search results	Showing how many items are retrieved	S5: She checked 226 items retrieved from the query of "coca cola advertisements."
Support users to examine different aspects of search results	Elements of surrogates/ meta-information; short description	S22: she read titles, descriptions, dates, and other information about retrieved items presented in the search result.
Support users to identify the format of an item.	Surrogates/ meta-information	S26: He identified the format of items from the information presented in search results.
Support users to judge the relevance or usefulness of an item retrieved.	Highlighting search terms; Surrogates/meta-information; short description/ Thumbnails	S36: Search terms were highlighted in the research results of historic newspaper archives.
Support users to understand the collections retrieved	Summary	S14: She read summaries of retrieved collections in search results.
Support users to extract relevant information from search results	Surrogates/meta-information; short description	S10: She found that President McKinley is one of the assassinated presidents from the search results.

4.2.1.3. Types of System Support for Mod Tactics

When users are unsatisfied with search results, they typically either modify the previous query or shift to another search strategy (Xie & Benoit 2013). Query modification was more frequently used than switching search strategies in this study. Users either specialize or broaden previous terms, or change terms with similar characteristics in an attempt to get more precise or expanded search results (Rieh & Xie, 2006).

System features for Mod tactics are overlapped somewhat with the ones for Creat tactics. For example, query suggestion can be also used for supporting users' query reformulation tactics. Users were able to select refined search terms from query suggestion. The most frequently used Mod method was search facets in search result pages. As digital resources are well structured with metadata in LOC-DL, different dimensions of facets were provided, such as format, date, source and language. Search facets were frequently used to limit their search results into a specific format or time range. However, Boolean operators were infrequently used for query reformulation.

Table 4-17. Types of system support for Mod tactics in LOC-DL

System support	Related system features	Example
Support users to find more relevant terms	Query expansion	S10: "jackie robinson" \Rightarrow "jackie robinson biography"
Support users to narrow down search results	Search facets; query expansion; advanced search; Boolean operators	S3: She limited the search results to "Online Format: Video" using the facet.
Support users to modify the format of search terms	Query expansion; query suggestion; spelling correction	S7: She corrected the spelling using query suggestion by the system ("assasinated" \Rightarrow "assassinated")
Support users to easily find, select and manipulate search facets	Location of facets; Forms of facets (e.g., facet checkboxes; drop-box selection of facets)	S3: She used search facet checkboxes located at the left side of the search results
Support users to come up with an appropriate modification strategy	Search suggestions	S10: She read the search suggestions offered by the system when encountering no results found.

4.2.1.4. Types of System Support for EvalI tactics

When a user enters into an individual item, she/he typically begins checking whether it is relevant or useful for the task. In Task 3, EvalI tactics were most frequently applied during the search session. Individual item evaluation requires a high degree of complexity and cognitive engagement, while representing primary value judgments (Xie & Benoit 2013). Table 4-18 summarizes types of system supports and corresponding system features that help users' application of EvalI tactics in LOC-DL. In order to help users judge relevance or usefulness of items effectively, LOC-DL provided key term highlighting, different elements of snippets, summary, short description, etc. Also, labels of sections or table of content were helpful for searchers to understand the structure of an item. When an item is long, subjects frequently scrolled down a page and grasped information needed by checking labels of subsections instead of reading all text from the top. Subjects also frequently used "Ctrl+F" to find specific information in an item, but it was the web-browser's feature, not LOC-DL's feature. Sometimes, subjects changed format or presentation of an accessed item. Zoom-in was a typical example. When viewing articles in newspaper collections, subjects enlarged PDF images of newspapers to make them easily legible. Similarly, the author observed that some subjects selected a text-format transcript when using audio interview files. They were able to quickly scan through the text transcript rather than listen to the entire interview audio.

Table 4-18. Types of system support for EvalI tactics in LOC-DL

System support	Related system features	Example
Support users to find key contents or concepts of the accessed item	Highlighting key terms; information snippets	S18: The item highlighted the search term "1964" using bold font.
Support users to have a brief idea about the accessed item	Summary/overview; short description	S36: She read the summary paragraph of an item.
Support users to understand the structure/layout of the accessed item	Table of content; labels for sections	S40: She used section labels to find the timeline of events related to assassination of President Lincoln.
Support users to understand the content of an item efficiently	Highlighting; information snippets; summary; descriptions	S43: Read the description of an item including title, summary, and notes.
Support users to select an alternative format of an item	Zooming; converted text of an scanned image	S17: She zoomed in a scanned text to easily read the content.
Support users to identify the format of an accessed item	File format description	S10: She checked the format information of an item.

4.2.1.5. Types of System Support for Xplor Tactics

In addition to search result evaluation, this study confirms that browsing strategies are frequently employed in using digital libraries. In particular, Xplor tactics were more frequently used than EvalR in exploratory search tasks.

As shown in Table 4-19, LOC-DL provided a series of system supports and associated system features for users' application of Xplor tactics. Like other digital libraries, LOC-DL also provided browsing categories by topic on the home page. Also, most individual collections showed some browsing categories in their main page. For example, American Memory Collections offered its own browsing categories with different criteria, such as

by topic, time period, format, and place. In most cases, there were short descriptions or subject terms briefly explaining about collections or items listed in the browsing categories. Basically, Xplor tactics required users' judgment of relevance or usefulness about individual collections or items listed in the predefined categories. Therefore, short descriptions, subject terms, or thumbnails were used as an aid for users judging the relevance of collections or items listed in categories. Users also frequently tried to find related items with an accessed item to further explore related topics by browsing through subject terms, related links, or references.

Table 4-19. Types of system supports for Xplor tactics in LOC-DL

System support	Related system features	Example
Support users to easily find categories for browsing	Location of browsing categories (homepage, main pages in each collections)	S57: She used "Browse Collections by Topic" in the main page of American Memory.
Support users to identify an appropriate criterion for browsing	Different browsing criteria (e.g., topic, time period, place, resource type)	S22: She browsed Coca-Cola advertising posters by year.
Support users to understand collections	Description on a collection; subject terms; labels; thumbnails	S55: He selected the category of "Coca-Cola Advertising" after reading the description.
Support users to identify related items with a current item	Subject terms; references; related links	S42: He selected a subject term to explore related items from the current item.
Support users to identify related resources with a particular collection	Related resources	S12: He surveyed related resources with the current collection.

4.2.1.6. Types of system support for AccF/AccB tactics

AccF tactics were used to access to specific items, browsing categories, or other pages during the search process in LOC-DL. AccF is one of the most frequently applied tactics

in all three types of search tasks. AccF tactics are a simple action, mostly clicking on a hyperlink to a new page or item. As AccF tactics are simple to apply compared to other search tactics, types of system support are also simple. To access to an item or other page, users selected hyperlinks. Like other common web pages, LOC-DL also differentiates hyperlinks from other regular text using different font style. Since all participants of this study have multiple years of experience in using the Internet, they did not encounter any special problems in relation to AccF tactics. Table 4-20 summarizes types of system support and related system features to assist users' application of AccF tactics.

Table 4-20. Types of system support for AccF tactics in LOC-DL

System support	Related system features	Example
Support users to find an access point to an item or other page	Different color, font (bold, underline, etc.); Thumbnails; Mouse hovering on links	S29: He clicked on the thumbnail of a Coca-Cola image that is linked to the item.
Support users to access to a full-text document	Link to a full text document	S57: She accessed a full-text file directly from the search result.

Users frequently applied AccB tactics to go back to previous search results or browsing categories in iterative loops. Also, they used AccB tactics to go to the homepage or the main page of an individual collection. The web browser's Back button was most frequently used to access back to previous pages. Since users were already familiarly with any web browser's back button, they liked to use the back button of a web browser rather than navigation bar or breadcrumbs provided by LOC-DL. When subjects opened multiple windows, they were able to go back to previous pages by closing the current

window. Table 4-21 summarizes types of system support and related system features for AccB tactics.

Table 4-21. Types of system support for AccB tactics in LOC-DL

System support	Related system features	Example
Support users to identify an access point to a previous item or page	Navigation bar; breadcrumbs; Hyperlinks to a main page	S50: He went to back to the main page of the collection of American Memory through the navigation bar.
Support users to go back to the starting point.	Hyperlinks to a main page	S30: He clicked the link to "Home" to go back to the main page of the American Memory collection.

4.2.2. Measuring User Perceptions of Search Support

This study intends to measure users' perceptions of search processes focusing on system support, difficulty, and satisfaction in terms of applying search tactics. After Task 3, subjects were asked to rate their perceived system support and their perceived difficulty in applying search tactics. The results of this section indicate which tactics were more or less supported by LOC-DL from the perspective of users, and which tactics were difficult to apply during the search process. In addition, users' perceived satisfaction levels were assessed for search tactic application.

4.2.2.1. Users' Perceived System Support

First, this study measured users' perceived system support for each type of search tactic using a five-point Likert scale (Table 4-22). The ANOVA test results reveal that there was a significant difference of perceived system support amongst different types of search tactics, $F(5.059, 298.506) = 5.637, p < 0.01$. Post-hoc pair-wise comparison tests

confirmed significant differences between AccF and Mod, AccF and EvalI, AccB and EvalI, and Xplor and EvalI at the alpha level of 0.05. The results showed that users experienced relatively high system support for application of AccF (3.50), AccB (3.45) and Xplor (3.43) tactics. AccF and AccB are relatively simple search tactics that require less cognitive and physical demand. Users perceived that system support would be sufficient for them to apply AccF and AccB tactics because these types of tactics were relatively easy to apply and accordingly users required less system support. Perceived system support for Xplor tactics also turned out comparatively high. LOC-DL provided well-structured browsing categories by topic or other criteria across different collections.

On the other hand, users' ratings of system support for EvalI (2.77) and Mod (2.80) turned out to be relatively low. EvalI tactics usually require a high degree of intellectual engagement (Xie & Benoit, 2013), but it seemed that LOC-DL provided insufficient system support to relieve users' intellectual loads required in evaluating activities. Also, the results revealed that users perceived relatively low system support for Mod tactics. Although LOC-DL offered different aspects of search facets in the search result page, there were limited system features of search term suggestions or controlled vocabularies.

Table 4-22. Users' perceived system support for each type of search tactic

Support for Tactics	Mean	STD	F (Greenhouse-Geisser)
Support for Creat	3.18	1.157	F(5.059, 298.506) = 5.637 p<0.01
Support for Mod	2.80	1.312	
Support for EvalR	3.13	1.142	
Support for AccF	3.50	1.157	
Support for AccB	3.45	1.156	
Support for EvalI	2.77	1.198	
Support for Xplor	3.43	1.254	

Table 4-23. Mean difference^a of pairwise comparisons^b between tactics

Tactics	Creat	Mod	EvalR	AccF	AccB	EvalI
Mod	.383					
EvalR	.050	.333				
AccF	.317	.700*	.367			
AccB	.267	.650	.317	.050		
EvalI	.417	.033	.367	.733*	.683*	
Xplor	.250	.633	.300	.067	.017	.667*

* p<0.05; ** p<0.01; ^a absolute value; ^b adjustment for multiple comparisons: Bonferroni

4.2.2.2. Users' Perceived Difficulty in Applying Search Tactics

Users' perceived difficulty was measured for each type of search tactic. Ratings of perceived difficulty showed the opposite results to the results from measuring perceived system support. Users responded that EvalI and Mod tactics were more difficult to apply while AccF and AccB tactics were easier. According to the ANOVA results, there were statistically significant differences of perceived difficulty levels between different types of search tactics, $F(4.668, 275.440) = 13.152$, $p < 0.01$. Table 4-24 presents post-hoc test results on pair-wise comparisons. Complexity of each type of search tactic would be closely related to users' perceived difficulty. Evaluating and query modification require more complicated user engagement, while accessing forward and backward are relatively simple actions. Users experienced less difficulty for application of Xplor tactics. Xplor tactics relied on predefined categories or lists, rather than users' creation of search statement. This implies a lesser amount of users' cognitive engagement is needed because users do not need to convert their search need to a specific form of statement in Xplor tactics. Instead of creating their own search statement, users can select some categories from the predefined list. In this sense, users are more passive in Xplor tactics compared

to query creation and result evaluation tactics. Query creation and search result evaluation, which build a "search result evaluation strategy" together, showed moderate difficulty, with ratings of 2.23 and 2.27 respectively. Even though query creation and result evaluation also involves a high level of user engagement, users who have been experienced in web searching for years seem to be familiar with Creat and EvalR tactic application. Interestingly, users perceived that EvalI would be more difficult to apply than EvalR, but the difference was not statistically significant.

Table 4-24. Users' perceived difficulty in applying each type of search tactic

Difficulty for Tactics	Mean	STD	F (Greenhouse-Geisser)
Difficulty for Creat	2.23	1.064	F(4.668, 275.440) = 13.152 p<0.01
Difficulty for Mod	2.53	1.096	
Difficulty for EvalR	2.27	1.118	
Difficulty for AccF	1.75	0.856	
Difficulty for AccB	1.82	0.948	
Difficulty for EvalI	2.58	1.157	
Difficulty for Xplor	1.93	0.861	

Table 4-25. Mean difference^a of pairwise comparisons^b between tactics

Tactics	Creat	Mod	EvalR	AccF	AccB	EvalI
Mod	.300					
EvalR	.033	.267				
AccF	.483*	.783**	.517**			
AccB	.417	.717**	.450*	.067		
EvalI	.350	.050	.317	.833**	.767**	
Xplor	.300	.600**	.333	.183	.117	.650**

* p<0.05; ** p<0.01; ^a absolute value; ^b adjustment for multiple comparisons: Bonferroni

This study further investigated how system support and difficulty would be related to each other. In all seven tactics, negative correlations were observed between system support and difficulty. Table 4-26 shows correlation analysis results. In all cases, Pearson r coefficients turned out statistically significant at the alpha level of 0.01. This result implies that more system support for search tactics would be associated with less difficulty in applying search tactics, and vice versa. In other words, users who experienced more difficulty in applying search tactics would perceive less system support, and vice versa.

Table 4-26. Correlation between system support and difficulty

Tactic	Pearson r between system support and difficulty
Creat	-.573 ^{**}
Mod	-.490 ^{**}
EvalR	-.520 ^{**}
AccF	-.590 ^{**}
AccB	-.480 ^{**}
EvalI	-.535 ^{**}
Xplor	-.365 ^{**}

* $p < 0.05$; ** $p < 0.01$

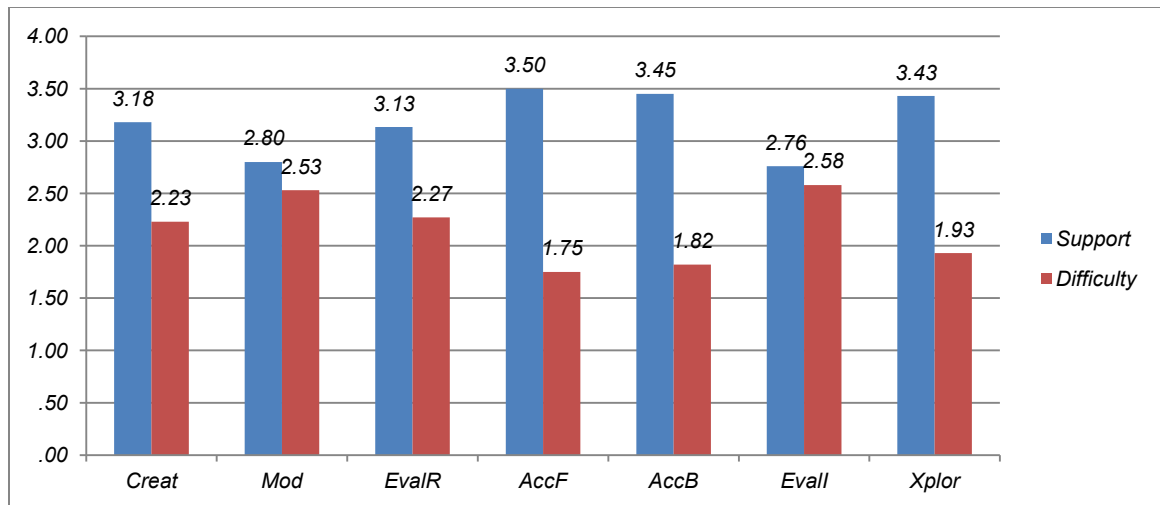


Figure 4-21. Comparison between perceived system support and difficulty levels in applying search tactics

4.2.3. Measuring Users' Perceived Satisfaction with Search Process

At the affective level, user satisfaction was measured for each type of search tactic using a five-point scale. Again, there were significant differences of satisfaction levels amongst different types of search tactics, $F(5.005, 295.312) = 9.493$, $p < 0.01$. The results of satisfaction measurement showed a similar pattern with perceived system support. Users' satisfaction ratings were relatively high in AccF (3.68), AccB (3.62), and Xplor (3.47). Again, AccF and AccB tactics do not involve high complexity of user engagement, so users experienced high satisfaction while applying these two types of tactics. Also, they were likely to experience relatively high satisfaction with the application of Xplor tactics.

On the contrary, users reported relatively low satisfaction levels with the application of Mod (3.05), EvalR (3.13), and EvalI (2.97) tactics. Higher complexity and less perceived system support were associated to users' lower satisfaction with Mod, EvalR and EvalI

tactics, and vice versa. In this respect, satisfaction would be closely correlated with perceived system support and less difficulty in search tactic application.

Table 4-27. User satisfaction in applying each type of search tactic

Satisfaction with search tactics	Mean	STD	F (Greenhouse-Geisser)
Satisfaction with Creat	3.23	1.110	F(5.005, 295.312) = 9.493 p<0.01
Satisfaction with Mod	3.05	.910	
Satisfaction with EvalR	3.13	1.081	
Satisfaction with AccF	3.68	.965	
Satisfaction with AccB	3.62	1.075	
Satisfaction with EvalI	2.97	1.041	
Satisfaction with Xplor	3.47	.895	

* p<0.05; ** p<0.01

Table 4-28. Mean difference^a of pairwise comparisons^b between tactics

Tactics	Creat	Mod	EvalR	AccF	AccB	EvalI
Mod	.183					
EvalR	.100	.083				
AccF	.450*	.633**	.550**			
AccB	.383	.567*	.483	.067		
EvalI	.267	.083	.167	.717**	.650**	
Xplor	.233	.417*	.333	.217	.150	.500**

* p<0.05; ** p<0.01; ^a absolute value; ^b adjustment for multiple comparisons: Bonferroni

This study also examined how satisfaction level would be associated with system support and difficulty in search tactic application. As shown in Table 4-29, there were positive relationships between satisfaction and system support, but negative relationships between satisfaction and difficulty. This reveals that users' perceptions of system support, difficulty and satisfaction would be all closely related with each other in the search process.

Table 4-29. Correlation between satisfaction, system support, and difficulty for each type of search tactic (Pearson r)

Tactic	System support	Difficulty
Satisfaction with Creat	.599**	-.561**
Satisfaction with Mod	.434**	-.474**
Satisfaction with EvalR	.548**	-.703**
Satisfaction with AccF	.448**	-.575**
Satisfaction with AccB	.632**	-.387**
Satisfaction with EvalI	.496**	-.516**
Satisfaction with Xplor	.241**	-.479**

* $p < 0.05$; ** $p < 0.01$

4.3. Effects of User Search Tactic Application on Search Outputs

The final results of this chapter answer the third research question, which examines the causal relationships between 1) user tactic selections and search outputs (e.g., search efficiency, success rate, aspectual recall, and satisfaction with search results) and 2) users' perceived system support and search outputs. Based on multiple regression, this study tried to identify which search tactics would influence search outputs, measured by efficiency, success rate, aspectual recall and satisfaction with search results.

4.3.1. Effects of Search Tactic Selections on Search Efficiency - Task 1

First, the effects of search tactic selections on search efficiency were examined with the data of Task 1. As mentioned in 4.1.1.1, fifty nine sessions in sixty were successfully completed by users finding the requested item, but session length differed by session. In this analysis, it was assumed that a shorter session is more efficient, and vice versa.

Session length has been used as one of the major variables to represent search session

efficiency in usability tests (Shackel, 1991; Nielson, 1993; Battleson et al, 2001; Joo, 2011). This study investigated the correlations between search tactic frequency and session length. As session lengths differed by session, standardized search tactic frequencies were used as explained in 4.1.1.4. Table 4-30 shows Pearson r coefficients between search tactic frequency and session length in different types of search tactics.

The frequency of query creation (Creat) shows a negative relationship with session length at the alpha value of 0.01 ($r = -0.617$). This reveals that users who applied Creat tactics more frequently would have more chances to complete their task quickly in Task 1. Also, the frequency of EvalI was negatively related to session length. In Task 1, EvalI was a relatively short tactic because it was simply to check whether the accessed item was the correct one requested by the assigned task. In Task 1, it took only 7.19 seconds for applying an EvalI tactic on average. Therefore, more uses of EvalI tactics resulted in shorter sessions in Task 1. More application of AccF tactics led to shorter sessions ($r = -0.308$; $p < 0.05$). EvalR showed a negative relationship (-0.223) but it was not statistically significant. Overall, application of query creation, accessing to an item, and item evaluation turned out to be more efficient way to complete the task quickly in known-item search tasks.

On the other hand, the frequency of AccB tactics was positively associated with session length ($r = 0.431$). As observed in 4.2.1.6, AccB tactics were typically used for iterative loops or going back to a previous page when failing to find a relevant item. Therefore, it usually takes a longer time when users walk through loops or go back to previous steps. In this respect, application of AccB tactics could lead to less efficient searches.

Table 4-30. Correlation coefficients between search tactic selection and search efficiency (Pearson r)

Search tactic frequency	Session length
Creat	-0.617**
Mod	0.065
EvalR	-0.223
AccF	-0.308*
EvalI	-0.528**
AccB	0.431**
Xplor	-0.070

* $p < 0.05$; ** $p < 0.01$

To investigate the causal relationship between search tactic frequency and session length, a multiple regression was conducted. When seven independent variables were entered into the regression model simultaneously, a significant violation of multicollinearity was observed in some variables, according to the criteria at 0.2 of Tolerance or 10 of VIF. In order to avoid collinearity problems, a stepwise method was employed instead. The stepwise method came up with a regression model with three predictors including frequencies of Creat, AccF, and AccB. An R^2 of 0.625 was achieved. Table 4-31 presents the regression analysis result with three predictors. In this regression model, the frequencies of Creat and AccF would negatively influence session lengths, while the frequency of AccB tactics would positively influence. This implies that query creation strategies would be more efficient in known-item searches. However, iterative sequences including AccB tactics might result in longer search sessions.

Table 4-31. Regression of session length on search tactic frequency in Task 1 (Stepwise method)

Predictor	B	Std. Error	β	t	Sig.
(Constant)	156.234	10.301			
Frequency of Creat	-51.925	7.703	-.562	-6.741	.000
Frequency of AccF	-15.359	3.320	-.382	-4.626	.000
Frequency of AccB	42.276	9.471	.375	4.464	.000

4.3.2. Effects of Search Tactic Selections on Task Success Rate - Task 2

For Task 2, the relationship between search tactic frequencies and success rates was investigated. Pearson correlations were computed between standardized search tactic frequencies and success rates in sixty search sessions of Task 2. Table 4-32 indicates correlative relationships between search tactic frequency and success rate in Task 2. Frequency of EvalI tactics turned out positively correlated with success rate by showing an r of 0.274. Users would get better search results when evaluating more individual items. Interestingly, frequency of Xplor tactics was negatively correlated with success rate in specific information search tasks by showing a Pearson r of -0.262 ($p < 0.05$). This implies that browsing strategies might not be an appropriate approach in achieving specific information searches.

Table 4-32. Correlation coefficients between search tactic selection and success rate (Pearson r)

Search tactic frequency	Success rate
Creat	0.053
Mod	0.011
EvalR	0.220
AccF	0.027
EvalI	0.274*
AccB	0.183
Xplor	-0.262*

* $p < 0.05$; ** $p < 0.01$

To further scrutinize causal relationships, a multiple regression was conducted. When entering seven predictors at the same time, there was a significant violation of multicollinearity in the regression model. Again, a stepwise method was selected as an alternative way to free from collinearity issues. Table 4-33 shows the regression analysis result based on stepwise method. Two predictors were identified in this model, which are frequencies of EvalI and Xplor. The frequency of EvalI would positively affect success rate in Task 2. It can be interpreted that individual item evaluation would be closely related to search outputs in a positive way. Selection of Xplor tactics has a negative impact on success rate. This suggests that browsing would not be an effective strategy in finding specific information in digital libraries.

Table 4-33. Regression of success rate on search tactic selection in Task 2 (Stepwise method)

Predictor	B	Std. Error	β	t	Sig.
(Constant)	.626	.060		10.500	
Frequency of EvalI	.160	.072	.271	2.212	.031
Frequency of Xplor	-.116	.055	-.259	-2.111	.039

The time data was also used to analyze the relationship between search tactic application and success rate. Correlation analysis was carried out to see which search tactics would be correlated with success rate in terms of spent time on applying search tactics. As shown in Table 4-34, time spent on EvalR and Xplor turned out to be significantly associated with success rate at the alpha level of 0.05. Time spent on EvalR was positively correlated with success rate ($r = 0.281$), whereas time spent on Xplor was negatively ($r = -0.312$). The more time users spent on EvalR, the higher success rate they achieved. In Task 2, users were able to find relevant specific information from the search result pages. Many users acquired relevant information from reviewing search result pages. In this way, users who spent more time on EvalR tactics had a better chance to get a higher success rate. However, browsing tactics turned out to be less effective to obtain a high success rate in specific information searches.

Table 4-34. Correlation coefficients between time spent on applying search tactics and success rate (Pearson r)

Time spent on search tactics	Success rate
Creat	0.178
Mod	-0.154
EvalR	0.281 [*]
EvalI	0.155
Xplor	-0.312 [*]

^{*} $p < 0.05$; ^{**} $p < 0.01$

A multiple regression was ran using a stepwise method due to the multicollinearity problem among independent variables. Three predictors were identified from the stepwise method – time spent on Xplor, Mod, and EvalR, while an R^2 of 0.272 was achieved. In the regression model, time spent on EvalR positively influenced success rate in specific information search tasks. As users could find relevant information from search result pages, EvalR tactics were useful for them to achieve the task.

Table 4-35. Regression of success rate on time spent on search tactics in Task 2 (Stepwise method)

Predictor	B	Std. Error	β	t	Sig.
(Constant)	.709	.068		10.359	.000
Time spent on EvalR	.264	.097	.365	2.714	.009
Time spent on Xplor	-.530	.206	-.325	-2.574	.013
Time spent on Mod	-.876	.263	-.440	-3.328	.002

4.3.3. Effects of Search Tactic Selections on Aspectual Recall and Satisfaction with Search Results - Task 3

In Task 3, this study investigated how search tactic selection would be associated with search outputs measured by aspectual recall and satisfaction. Correlation analysis results reveal that aspectual recall rate would be associated with frequencies of AccF, EvalI, Xplor, and Obt tactics respectively. It is not a surprise that the frequency of Obt tactic application is closely related to aspectual recall rate ($r=0.687$) because item obtaining activities would directly increase the recall rate. Also, AccF and EvalI turned out to be positively related to aspectual recall. Interestingly, Xplor was moderately correlated with aspectual recall while query related tactics (Creat, Mod, and EvalR) were not. This implies that users would be able to obtain more useful items from Xplor tactics rather than query-related tactics in exploratory search tasks.

In addition, this study examined the relationship between search tactic selections and satisfaction with search results. Frequencies of EvalI and Obt tactics are significantly correlated to satisfaction with search results. EvalI and Obt tactics led directly to finding relevant information, so users would feel more satisfactory with search results from application of these two types of search tactics.

Table 4-36. Correlation coefficients between search tactic selection and search outputs (Pearson r)

Tactic	Aspectual recall	Satisfaction with search results
Frequency of Creat	-0.237	-0.112
Frequency of Mod	0.226	-0.171
Frequency of EvalR	-0.237	-0.131
Frequency of AccF	0.479**	0.154
Frequency of EvalI	0.580**	0.258*
Frequency of AccB	0.191	-0.094
Frequency of Xplor	0.352**	0.054
Frequency of Obt	0.687**	0.390**

* $p < 0.05$; ** $p < 0.01$

To examine the causal relationship between search tactic application and search outputs, a multiple regression was conducted. Obt tactic was excluded in the regression model, as it was an obvious predictor to explain aspectual recall rate. Table 4-37 provides the multiple regression result with seven independent variables ($R^2=0.480$). However, multicollinearity existed amongst predictors according to the collinearity diagnosis criteria, which are set as 0.2 of Tolerance and 5 of VIF.

To avoid multicollinearity between independent variables, a stepwise method was used in regression analysis. Five independent variables were eliminated from stepwise entering method. The result generated an estimation model of aspectual recall explained by the frequencies of EvalI and Xplor tactics. An R^2 value of 0.404 was achieved. This result reveals that the frequencies of EvalI and Xplor tactics would affect aspectual recall rate. The more EvalI and Xplor tactics users applied, the higher aspectual recall they achieved in exploratory searches.

Table 4-37. Regression of aspectual recall on frequencies of EvalI and Xplor tactics (Stepwise method)

Predictor	B	Std. Error	β	t	Sig.
(Constant)	0.074	0.023		3.165	0.002
Frequency of EvalI	0.007	0.001	0.537	5.222	0.000
Frequency of Xplor	0.005	0.002	0.263	2.562	0.013

A multiple regression was carried out between search tactic selections and satisfaction with search results. Since collinearity was detected among predictors, a stepwise method was employed again in the analysis. Six predictors were removed from the stepwise method, which means only one significant predictor remained to account for satisfaction with search results. A regression model of satisfaction with search results on EvalI tactic frequency was identified. However, an R^2 value was too low to adequately account for satisfaction with search results ($R^2=0.066$). This result reveals that frequency of EvalI tactics would affect users' satisfaction with search results. However, only about 6.6 percent of the variance in satisfaction level with search results were explained by the frequency of EvalI tactics.

Table 4-38. Regression of satisfaction with search results on frequency of EvalI tactics (Stepwise method)

Predictor	B	Std. Error	β	t	Sig.
(Constant)	2.706	.362		7.483	.000
Frequency of EvalI	.048	.024	.258	2.031	.047

In addition, the relationship between search tactic application and search outputs were analyzed in terms of time spent on tactics. As shown in Table 4-39, correlation

coefficients were calculated between time spent on each tactic and aspectual recall and satisfaction with search results. Aspectual recall turned out to be significantly correlated with time spent on EvalR, EvalI, Xplor, and Obt tactics. Interestingly, aspectual recall was negatively related to time spent on EvalR. Longer time spent on search result evaluation did not necessarily lead to higher aspectual recall. This reveals that the time spent on EvalR tactics might not be an effective approach to collecting different aspects of information on a topic in exploratory searches. Time spent on EvalI, Xplor, and Obt tactics would be positively associated with aspectual recall. The more time users spent on browsing, individual item evaluation, and obtaining items, the higher the aspectual recall was achieved. As to satisfaction with search results, only time spent on obtaining tactics turned out to be significantly related at the alpha value of 0.05. Obviously, obtaining tactics would have some direct connection to search results, and accordingly more time spent on Obt tactics would be linked with users' satisfaction level with search results.

Table 4-39. Correlation coefficients between time spent on search tactics and search outputs (Pearson r)

Tactic	Aspectual recall	Satisfaction with search results
Time spent on Creat	-0.211	-0.044
Time spent on Mod	-0.148	-0.075
Time spent on EvalR	-0.294 [*]	0.011
Time spent on EvalI	0.281 [*]	0.029
Time spent on Xplor	0.278 [*]	-0.086
Time spent on Obt	0.323 [*]	0.287 [*]

^{*} $p < 0.05$; ^{**} $p < 0.01$

A multiple regression was run with time data of search tactic application and aspectual recall. Again, the variable of time spent on Obt tactics was excluded. Table 4-40 presents

a regression analysis result with collinearity tests. The indices of collinearity, including Tolerance and VIF, verified that there is no serious interrelationship between predictors. The regression model attained an R^2 of 0.233. In this regression model, regression weights of EvalI and Xplor were statistically significant at the alpha level of 0.05, which are 0.319 and 0.301 respectively. It is concluded that time spent on EvalI and Xplor tactics would affect aspectual recall rate in exploratory search tasks. Interestingly, this result is consistent with the findings from the frequency data analysis shown in Table 4-32.

Table 4-40: Regression of aspectual recall on time spent on search tactics

Predictor	B	Std. Error	β	t	Sig.
(Constant)	.123	.048		2.540	.014
Creat	-.001	.001	-.170	-1.354	.181
EvalI	.000	.000	.319	2.499	.016
EvalR	.000	.000	-.089	-.541	.591
Mod	.000	.001	.021	.145	.885
Xplor	.001	.000	.301	2.302	.025

A multiple regression was conducted with the dependent variable of satisfaction with search results. As shown in Table 4-41, there was no predictor that has a significant impact on users' satisfaction with search results. An R^2 value of the model was also very low, which is 0.019. This suggests that time spent on search tactics would not influence users' satisfaction level in terms of search results.

Table 4-41: Regression of time spent on search tactics on satisfaction with search results

Predictor	B	Std. Error	β	t	Sig.
(Constant)	3.202	.688		4.652	.000
Creat	-.005	.014	-.053	-.371	.712
EvalI	.000	.002	.012	.086	.932
EvalR	.001	.003	.075	.404	.688
Mod	-.007	.010	-.118	-.709	.481
Xplor	-.002	.004	-.078	-.527	.600

4.3.4. Effects of Search Processes on Search Outputs

In order to examine how system support would be associated with search outputs, this study calculated correlation between users' perceived system support and aspectual recall and satisfaction with search results. First, EvalR, AccF, EvalI and Xplor turned out to be significantly related with aspectual recall. In particular, high positive correlation was observed between system support for Xplor and aspectual recall ($r = .581$). Also, system support for two evaluation tactics, EvalR and EvalI, showed moderate high correlation with aspectual recall, 0.385 and 0.482 respectively, at the alpha level of 0.05. However, system support for query creation and reformulation tactics would not be significantly associated with aspectual recall according to the correlation analysis.

Users' perceived system support was more closely related to their satisfaction level of search results. Except Mod, system support for all types of search tactics investigated were related to satisfaction with search results. Correlation between support for EvalI and satisfaction with search results turned out to be highest (0.555). Also, system support for

EvalR, AccF, and Xplor would be moderately related to satisfaction level in terms of search results at the alpha level of 0.01.

Table 4-42. Correlation coefficients between the perceived system support and search outputs (Pearson r)

Perceived Support	Aspectual recall	Satisfaction with search results
Support for Creat	-0.070	.328*
Support for Mod	0.164	.194
Support for EvalR	0.385**	.442**
Support for AccF	0.317*	.457**
Support for AccB	0.128	.260*
Support for EvalI	0.482**	.555**
Support for Xplor	0.581**	.453**

* $p < 0.05$; ** $p < 0.01$

To examine the effect of system support on search outputs, a multiple regression analysis was performed. The aspectual recall was regressed on system support for search tactics. Table 4-43 presents a regression analysis result with seven independent variables in the model. There was no problematic collinear relationship among independent variables. An R^2 value of 0.420 was achieved in the model. The result indicates that system support for Xplor tactics affects aspectual recall rate at the alpha level of 0.01, and a standardized regression loading of 0.441 was observed. Interestingly, according to the regression coefficients computed, system support for Creat would negatively influence aspectual recall ($\beta = -0.295$). This result reveals that system support for browsing features would be more important than query creation features to attain higher aspectual recall.

Table 4-43. Regression of perceived system support on aspectual recall

Predictor	B	Std. Error	β	t	Sig.
(Constant)	.079	.037		2.114	.039
Support for Creat	-.021	.008	-.295	-2.439	.018
Support for Mod	.005	.007	.084	.731	.468
Support for EvalR	.011	.011	.150	.997	.324
Support for AccF	.002	.009	.025	.206	.837
Support for AccB	-.002	.008	-.025	-.235	.815
Support for EvalI	.019	.010	.282	1.966	.055
Support for Xplor	.029	.007	.441	4.036	.000

Dependent variable: aspectual recall

Next, the effect of system support on satisfaction with search results were examined based on regression analysis. An R^2 of 0.436 was achieved when entering seven predictors. Regression coefficients of system support for EvalI and Xplor turned out to be statistically significant at the alpha level of 0.05. The obtained standardized regression weights were 0.359 and 0.279 respectively for EvalI and Xplor. This reveals that system support for these two tactics would positively affect satisfaction level to search results. In exploratory search tasks, users' perceptions of system support for EvalI and Xplor are important to increase users' satisfaction with search results.

Table 4-44. Regression of satisfaction with search tactics on aspectual recall

Predictor	B	Std. Error	β	t	Sig.
(Constant)	.690	.533		1.295	.201
Support for Creat	.189	.121	.197	1.553	.126
Support for Mod	-.050	.102	-.059	-.487	.628
Support for EvalR	-.039	.154	-.041	-.255	.799
Support for AccF	.143	.124	.150	1.156	.253
Support for AccB	.027	.108	.028	.253	.802
Support for EvalI	.331	.139	.359	2.381	.021
Support for Xplor	.246	.101	.279	2.431	.019

Dependent variable: satisfaction with search results

4.3.5. Structural Path Model of Search Process on Search Output

In order to explain the relationships between variables more comprehensively, a structural equation modeling (SEM) analysis was carried out. The model to be examined is designed as a diagram in Figure 4-22. Basically, this model was constructed to investigate how "users' search tactic selections", "system support for search tactics", and "satisfaction with search tactic application" would affect "aspectual recall" and "satisfaction with search results". In order to obtain the variables of "satisfaction with search tactics" and "system support for search tactics" in a collective way, the author set a unit of analysis as each type of search tactic in a session. As seven types of tactics were investigated in a session, 420 observations were made for each variable.

The proposed structural model contains the following variables:

- Observed, endogenous variables: satisfaction with search process (Y_1), aspectual recall (Y_2), satisfaction with search results (Y_3)

- Observed, exogenous variables: search tactic selection - tactic frequency (X_1), system support for search tactics (X_2), and difficulty for search tactics (X_3)
- Unobserved, disturbances: D_1 , D_2 , and D_3

The SEM model is specified as follows:

- $Y_1 = \gamma_{12}X_2 + \gamma_{13}X_3 + D_1$
- $Y_2 = \gamma_{21}X_1 + \gamma_{22}X_2 + \beta_{21}Y_1 + D_2$
- $Y_3 = \gamma_{31}X_1 + \gamma_{32}X_2 + \beta_{31}Y_1 + \beta_{32}Y_2 + D_3$
- $$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 0 & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & 0 \\ \gamma_{31} & \gamma_{32} & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} + \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix}$$
- $$\underline{\mathbf{Y}} = \mathbf{\Gamma} \underline{\mathbf{X}} + \mathbf{B} \underline{\mathbf{Y}} + \underline{\mathbf{D}}$$

where Γ matrix indicates directional relationships from exogenous variables to endogenous variables; B matrix indicates directional relationships from endogenous variables to endogenous variables.

Also, variance/covariance matrices are specified as follows:

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \quad \Psi = \begin{bmatrix} \psi_{11} & 0 & 0 \\ 0 & \psi_{22} & 0 \\ 0 & 0 & \psi_{33} \end{bmatrix}$$

where Φ indicates the covariance matrix of exogenous variables (symmetric); Ψ indicates the variance matrix of the disturbances (symmetric).

The established structural path model includes eighteen parameters to be estimated, which are 9 regression coefficients (Γ matrix and B matrix), 3 covariances and 3 variances of exogenous variables (Φ matrix), and 3 variances of disturbances (Ψ matrix). The model was fitted with the observed data based on maximum likelihood. The SEM result achieved an adequate model fit: RMR=.039, GFI=.989, AGFI=.925, NFI=.976, and CFI=.980. Figure 4-22 presents parameters estimated, including standardized regression coefficients and correlation coefficients.

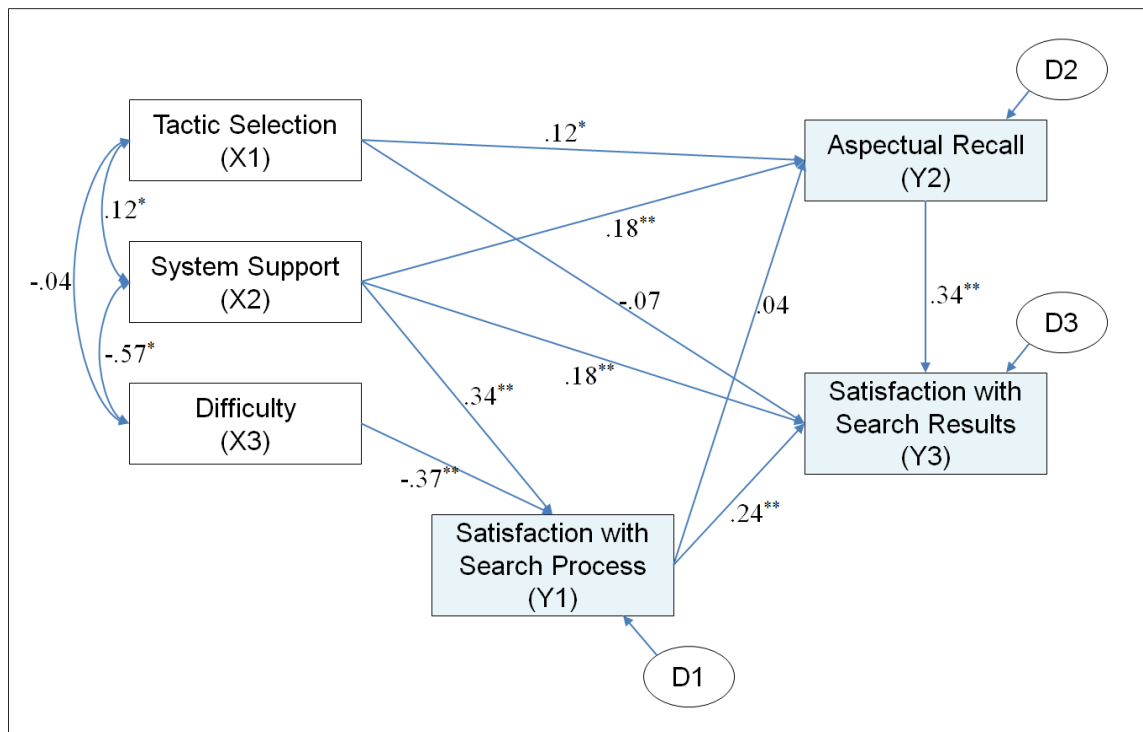


Figure 4-22. A structural path model (* $p < .05$; ** $p < .01$)

The SEM result reveals causal relationships between exogenous and endogenous variables. This model well summarizes overall relationships between variables that we investigated in the previous sections. Table 4-45 shows estimated parameters of regression weights. First, "search tactic selection" would affect "aspectual recall", but not "satisfaction with search results". Second, "system support for search tactics" would

affect both "aspectual recall" and "satisfaction with search results". Third, "satisfaction with search tactic application" would affect "satisfaction with search results", but not "aspectual recall". Fourth, "aspectual recall" would affect "satisfaction with search results".

In addition, the author tried to insert "difficulty for search tactic application" into the model. This model assumes that "difficulty" has an indirect effect to search outputs through "satisfaction with search tactics." The effects of "system support" and "difficulty" on "satisfaction with search tactics" were examined. "System support" influences "satisfaction with search tactics" positively by showing a regression coefficient of 0.34 at the alpha level of 0.01. On the contrary, "difficulty" exhibited a negative effect on "satisfaction with search tactics". This affirms an obvious finding that users are likely to feel less satisfactory when they experience difficulty in applying search tactics.

Table 4-45. Regression weights estimated in the structural path model

Parameter	Directional relationship	Standard error	Estimate ^a	Critical Ratio	P value
γ_{12}	System support \Rightarrow Satisfaction with search process	.042	.335	7.234	***
γ_{13}	Difficulty \Rightarrow Satisfaction with search process	.045	-.371	-8.003	***
γ_{21}	Tactic selection \Rightarrow Aspectual recall	.001	.121	2.548	.011*
γ_{22}	System support \Rightarrow Aspectual recall	.004	.180	3.169	.002**
γ_{31}	Tactic selection \Rightarrow Satisfaction with search results	.007	-.067	1.604	.109
γ_{32}	System support \Rightarrow Satisfaction with search results	.047	.176	3.530	***
β_{21}	Satisfaction with search tactics \Rightarrow Aspectual recall	.004	.038	.667	.505
β_{31}	Satisfaction with search process \Rightarrow Satisfaction with search result	.051	.242	4.951	***
β_{32}	Aspectual recall \Rightarrow Satisfaction with search results	.576	.343	8.110	***

* p<.05; ** p<.01; *** p<.001; ^a standardized

In addition, parameters of covariances between exogenous variables were estimated in the SEM model. Table 4-46 presents estimated coefficients of correlations. There was negative association between "system support" and "difficulty". Interestingly, there was a significant relationship between "search tactic selection" and "system support" ($r = .122$) at the alpha level of .05. This suggests that users would perceive more system support for a certain type of search tactic when they apply them more frequently. There was no correlation found between "search tactic selection" and "difficulty".

Table 4-46. Correlations between the exogenous variables

Parameter	Relationship	Covariance estimate	Standard error	Correlation ^a	Critical Ratio	P value
Φ_{12}	Tactic selection \Leftrightarrow System support	.887	.359	.122	2.472	.013*
Φ_{13}	Tactic selection \Leftrightarrow Difficulty	-.268	.327	-.040	-.819	.413
Φ_{23}	System support \Leftrightarrow Difficulty	-.714	.070	-.570	-10.141	***

* $p < .05$; ** $p < .01$; *** $p < .001$; ^a standardized

4.4. Chapter Summary

This chapter answered three research questions. First, users' search tactic patterns were analyzed in terms of frequency, spent time, change over time, and transition. This study investigated unique patterns of search tactic selection in different search tasks in terms of both frequency and time. Changes in search tactic selection within a session were traced

based on non-parametrical analysis. Transitions in search tactics were explored to identify most frequent paths of user search tactic selection. Second, types of system supports were identified based on open coding. A range of system support types were identified for each type of search tactic in LOC-DL. Also, user perceptions of system support, difficulty, and satisfaction were measured for search tactic application. Users' perceived system support was relatively high for AccF, AccB, and Xplor tactics, but relatively low for EvalI and Mod tactics. Users also perceived high difficulty for application of EvalI and Mod tactics while relatively low difficulty for AccF and AccB tactics. Third, the effects of search tactic selections and perceived system support on search outputs were examined. In Task 1, frequencies of Creat and AccF tactics positively influenced search efficiency. In Task 2, frequency of EvalI had a positive impact on success rate whereas frequency of Xplor did a negative impact. Time spent on EvalR influenced success rate positively while time spent on Xplor and Mod did negatively. In Task 3, frequencies of EvalI and Xplor tactics and time spent on EvalI tactics positively affected aspectual recall. SEM analysis showed comprehensive relationships between search process variables and search output variables.

CHAPTER 5. DISCUSSION

A significant contribution of this study lies in 1) exploring users' search tactic selections and patterns in using a digital library; 2) evaluating search process by measuring system support, difficulty, and satisfaction with regard to the application of search tactics; and 3) examining effects of user search tactic selections on search outputs. In this chapter, the author discusses users' unique search tactic patterns and associated system support in digital library environments. Implications of IR system designs are also discussed based on the findings of this study. Additionally, some methodological implications are reviewed in relation to interactive IR evaluation.

5.1. Users' Unique Search Patterns and IR System Design Implications

In this section, the author reviews users' unique search behavior in digital libraries, and discusses IR system design implications to better support users' interactions with IR systems in the context of digital libraries.

5.1.1. Effects of Search Task Type on Search Tactic Selections

The findings of this study confirm that users' search tactic application patterns differ by task types. That is, search task type can be a factor that influences users' search tactic selections. This reaffirms previous studies that proved significant task effects on information search behavior (Vakkari, 2003; Hung 2005; Liu et al., 2010; Arguello et al., 2012; etc.).

First, in Task 1, users frequently applied a single query strategy, which indicates completing a search task with only one time query effort and consequent result evaluation. Thus, the proportion of Creat tactics showed relatively higher compared to the other two search tasks. Since users have prior information on the source to be retrieved, they tend to create a search query using their own knowledge about the item, such as title, subject, or time range. Prior knowledge about the item help users come up with a query more precisely and adequately, users are able to find the item quickly without modification of queries. Therefore, known-item search could be relatively easier task in terms of query creation while less cognitive loads are demanded in formulating queries since users already have some clues about the item to be searched. Since users are likely to finish their known-item task with a query creation with less modification, AccB and Mod tactics are less frequently selected. If the retrieval mechanism works well in an IR system, users would be able to complete this type of task straightforwardly and quickly. In this study, fifty nine out of sixty subjects successfully found the requested item within five minutes. Also, the average number of tactics applied in Task 1 was much less than the ones in the other two tasks. Xplor tactics can be also applied in known-item search tasks. Based on their prior knowledge, users can select proper collections from collection categories. In this study, browsing strategies showed longer paths than query creation approaches because users needed to get to deeper levels of the site to reach relevant items. Although it usually takes longer time, browsing can be an alternative strategy to attain known items without effort to create a query.

This suggests several implications for IR system design to support users' more efficient and effective search activities in known-item searches. Since users have already some

prior knowledge of the item such as title and creator, they usually create queries including those information, which are related to metadata. Well-structured and complete document representations, including various metadata elements of title and creator, are important to lead users to gain successful search results (Ogilvie & Callan, 2003).

Advanced search functions that incorporate structured data enable users to produce more precise queries especially in known-item search tasks (Yasunaga et al., 2013). In addition, federated search functions are useful in digital library environments. As most digital libraries consist of subordinate collections, it is important to empower users to search multiple collections at one attempt in known-item searches. Again, advanced search function based on structured metadata is the key to support users to accomplish known-item search, and a federated search function is also useful to efficiently search multiple collections in a digital library system.

Second, in specific information searches, users prefer to engage iterative search result evaluation. This study empirically observed that proportions of EvalR, AccF, and AccB were relatively high in specific information search sessions. Usually, information objects requested in specific information searches are pieces of information snippets, such as specific names, dates, or events, rather than thorough information on a particular subject. Information snippets can be usually obtainable from retrieval results, even not reviewing individual items. Users obtain relevant information snippets from surrogates or meta-information presented on search result pages. Hence, in support of specific information searches, how to properly present search results to enable users to collection relevant information snippets directly from the list of search results is important (Rose & Levinson, 2004; Cutrell & Guan, 2007). Thus, well-organized, rich information search

results could better support users to perform specific information searches. In addition, it would be useful if a digital library system also had a function to present quick answers instead of documents or full items to respond to the specific information search tasks based on structured queries (Mika, 2008). Thus, summarized information about an item could help users quickly find specific information, instead of full-text document. In addition, highlighting key information would be another support for users to fetch specific information.

Third, it is widely acknowledged that exploratory tasks require more support as search process involve more interactions in exploratory searches (Diriye, Blandford, & Tombros, 2010; Kules & Carpa, 2011). In this study's Task 3, users spent the longest time on evaluating individual items. Basically, to accomplish an exploratory search task, users have to visit different pages and items to collect different aspects of information on a particular subject. Unlike known-item searches or specific information searches, users usually require more in-depth, comprehensive information about a particular topic in exploratory searches. Thus, EvalI tactics play a key role in accomplishing an exploratory search task. Accordingly, system support for EvalI tactics is essential to support this type of search task in designing digital library systems. System design implications related to supporting EvalI tactics are discussed in detail in Section 5.1.3.

Xplor is another frequently applied tactic in exploratory search tasks. This study also found that Xplor tactics would be effective in achieving better search outputs measured by aspectual recall. Browsing is an effective search strategy when users need to obtain a great deal of contextual information on a certain topic (Shen et al., 2006). As browsing is

frequently selected in exploratory searches, digital library systems need to facilitate users to better explore different categories, subject lists, item lists, or resource lists. This finding implies that digital library systems should focus on the enhancement of browsing functions for exploratory searches, such as various options for browsing, different browsing criteria (e.g., topic, date, region, etc.), and offering task-oriented browsing in addition to subject browsing (Mu, Ryu, & Lu, 2011). Also, since users often select a browsing strategy to start their exploratory session, the entrance of digital libraries should include proper browsing options to help users easily initiate a browsing method.

5.1.2. Search Tactic Selection Changes in a Single Session

Using kernel regression, this study investigated how users' search tactic patterns change over time within a single session. A search session usually begins with a query creation or browsing method in digital libraries. That is, users typically initiate their search session by forming a query or selecting a collection from categories when searching in a digital library system. Accordingly, this study confirms that Creat and Xplor are two dominant search tactics in the beginning phase of a search session. To be more specific, this study found that Creat tactics were more frequently used in the beginning of known-item searches (Task 1) or specific information searches (Task 2), while Xplor tactics were more often selected in exploratory searches (Task 3). Thus, in order to support users' starting of a session, it is essential to assist both query creation and browsing strategies. For the starting with a query, every suggestion concerning search function design could be applicable to help users initiate their session, such as a basic search box, advanced

search function, query suggestions, query error correction, and others. Particularly, query expansion would be useful to help users come up with an initial query to start a session. As to starting with a browsing strategy, topic categories presented on the homepage are crucial to lead users to successfully initiate their search session. Well-organized categories are important in support of users browsing strategies at the starting point of a session. Also, thumbnail display and dropdown menus are useful to facilitate users to navigate different topics of collections (Kang & Shneiderman, 2000). Presenting featured collections is also another feature to get attention from users onto a particular collection, which will facilitate users to start their search with a browsing strategy.

In the middle phase, probabilities of EvalR, Xplor or EvalI are relatively high. Users tend to engage in a search process by evaluating search results, browsing different categories or items, or evaluating individual items iteratively. For example, users repeatedly evaluated search results in Task 2, while they iteratively evaluate a series of individual items in Task 3. The results of this study showed that nearly 50% probability of EvalR was observed in the middle phase of session in Task 2. Similarly, EvalI tactics occurred with about a 50% chance in the middle phase of Task 3. In Task 2 and Task 3, iterative loops were users' main behavior in the middle phase of a search session. Thus, digital library systems should consider how to reduce unnecessary iterations involving search result evaluation (EvalR) or browsing (Xplor). Design implications about reducing iterations are discussed later in detail in Section 6.1.3.

In the ending phase, users typically close their search sessions by continuing patterns of the middle phase or engaging more evaluating activities. In this study, it was observed

that users' search tactic patterns remained stable from the middle phase to the end of the session in Task 3. On the contrary, in Task 1 and Task 2, users applied more EvalI or EvalR tactics before finishing the session. They tried to conclude a search session with a last attempt to find more relevant information. However, because of the time limit imposed on each session, this study could not investigate how users naturally finish their search session. Therefore, the design implications of the ending phase could not be discussed in this study.

5.1.3. Frequently Applied Search Strategies

From the analysis of search tactic transitions, this study identified two most frequently used search strategies, 1) iterative search result evaluation and 2) iterative browsing. These two search strategies are commonly applied in Web environments (Shen et al., 2006; Xie & Joo, 2010b; Zhang et al., 2012), and this study confirmed that these two strategies are also frequently applied in digital library searches. This suggests users' search tactic patterns are quite predictable as their transitions showed explicit patterns rather than random transitions. Wildemuth's (2004) study results that a few sequential combinations of moves are most frequently used in search tasks are also confirmed in digital library searches. Additionally, Olah's (2005) findings that a series of iterative loops constitutes an interaction process are reaffirmed by this study.

First, iterative search result evaluation showed repeated transition patterns among EvalR, AccF, EvalI, and AccB. This transition pattern indicates that a hub-and-spoke model (Catledge & Pitkow, 1995; Tidwell, 2011) can also be applied to digital library searches.

In this case, a hub-and-spoke model shows iterative returns to the search result page as to proceed to several items during the search process. To design more efficient iterative search result evaluation, it might need for digital library system designers to adopt an interface that presents search results and documents together simultaneously in the same screen. By showing search results and documents together at the same time, users can reduce unnecessary repetition of AccF and AccB tactics, which will result in shorter search paths. Previous researchers empirically proved that showing search results and an item together support users' efficient navigation in search process (Diriye et al., 2010; Mu et al., 2011; Golovchinsky et al., 2012). In addition, various system features can be applicable in digital library system design to support an iterative result evaluation strategy, such as well-structured meta information, categorization of search results, search result visualization, ranking based on user feedback, and others (Agichtein, Brill, & Dumais, 2006; Ahn & Brusilovsky, 2009; Zhang, 2008; Mu et al., 2011; Marchionini, 2006).

Second, browsing is another representative search strategy that is frequently selected in searching digital libraries. In particular, it was observed that iterative browsing was frequently applied in Task 3. Iterative browsing refers to a search strategy in which users browse and evaluate a series of items based on predefined topic categories or item lists (Xie & Joo, 2010b). Basically, information architecture is important in designing digital libraries to facilitate users' browsing tactics. Adequate organization of topic categories is essential to encourage users to browse various collections or items while searching in the digital library system. In this study, well-organized categories led to users' application of browsing strategy as shown in the following quote: "They (LOC-DL) are very nicely

organized ... you can click on different things (S40)... {then, started a browsing strategy}." The key to effective browsing is well-organized, flexible, and dynamic information presentation to end users (Kent & Bowman, 1995). This implies that digital libraries need to be equipped with functions for supporting a browsing strategy, such as well structured categories of sub-collections or items on a specific topic. Either browsing categories or faceted categories could be useful for users to explore information resources (Hearst, 2006). Also, taxonomy or classification scheme is needed to arrange a set of objects into categories with shared characteristics (Kent & Bowman, 1995). In exploratory searches, information foraging design can be useful to help users find different aspects of information on a specific topic. Exploratory search tasks are closely related to information foraging and sensemaking process. To facilitate users effective browsing activities, digital libraries should support users' information foraging by predicting users' navigation paths and designing foraging cues. Common techniques that support information foraging can be also applicable into digital library system design, such as listing related items, recommendations, and adaptive navigation (Olston & Chi, 2003; Brusilovsky et al., 2004; Piorkowski et al., 2012).

5.1.4. Infrequently Applied Search Tactics

In this study, it was found that minor tactics, including Lrn, Mon, Org, and Rec, were rarely used in digital library searches. This result implies the principle of least effort in users' information seeking. Users are inclined to engage in least effort activities during a search process (Bates, 2002). Rather than actively engaging in the search process by

applying different search tactics, users are likely to minimize their procedural effort in the search process. Since each search session was quite short in this study, no Mon and Rec tactics was observed. There were only a few of Lrn and Org tactics observed, but these two tactics comprised less than 1% of the entire tactics applied. More interestingly, users rarely seek explicit help with intent, such as search instructions or Help pages when they faced with a problem. Lrn is one of tactics where users actively gain knowledge about search skills and systems, such as how to operate the system, and it could lead to effective and efficient search processes (Xie & Cool, 2009). In spite of benefits of Lrn tactics, users infrequently used explicit help features in this study. Instead, trial and error was users' preferred approach to resolve problems they encountered. It is because of users' search preference, lack of credibility or usability, or unawareness of Help pages. For example, some subjects liked to solve a problem by themselves instead of referring to help functions as shown in the following quote: "I generally prefer to try things myself until it works (S9)." Some subjects believed that help pages are less useful or they just did not like to use help functions: "... .. this is from experience having dealt with a lot of computer systems, the help is very ill documented. It is usually too wordy (S53)." Org tactics were limitedly used in sorting search results. Users did not frequently sort any list of search results even though LOC-DL provided sorting options by different criteria. For browsing categories, there was no option to sort categories. In many digital library systems, browsing categories are usually fixed and predefined while not allowing users to change its presentation. With regard to Mon and Rec, search tasks of this study were too short to observe those tactics. Also, the study investigated only single sessions, not multiple sessions. Monitoring and recording tactics are more needed in multiple search

sessions or multitasking (Lin & Belkin, 2004; Lin, 2005; Du & Spink 2011) rather than single sessions. As this study concerned single sessions within restricted time on task, the author did not observe any Mon or Rec tactic.

5.1.5. Effects of Search Process on Search Output

This study is one of the few studies that examined causal relationships between IR process variables (e.g., frequency of search tactics, time spent on search tactics, satisfaction with search tactic application) and IR outputs (e.g., efficiency, success rate, aspectual recall, satisfaction with search results). Based on multiple regression and SEM, this study investigated how search tactic application would affect search outputs in different task situations.

In Task 1, it was found that frequent uses of Creat and AccF tactics led to shorter session length. Typical short sessions of Task 1 involved only one query creation during an entire session. When using a single query without any modification, users could finish the task quite quickly. In known-item search tasks, the shortest path to an item would be initial query creation and following evaluation. However, to successfully complete the task using only one query, the initial search terms must be relevant enough to bring relevant results at one time. Of course, query modification makes a session less efficient when a user fails to complete the task with one query. The findings of this study also suggest that searching would be more efficient than browsing in known-item search tasks. In addition, this study showed that frequent application of AccB tactics resulted in longer session

length, which means less efficiency. AccB tactics were usually used to go back to previous search results or categories when an accessed item was not relevant.

In Task 2, the frequency of EvalI and the time spent on EvalR would positively influence success rate. This study found that users prefer the iterative result evaluation strategy in specific information search tasks. Also, it turned out the iterative result evaluation strategy would be more effective to retrieve specific information in digital libraries. It was often observed that users found relevant information from surrogates of search results. In this sense, it is important to present appropriate surrogates of an item in search results to better support specific information searches. System features for iterative result evaluation will be useful to support users to search specific information from digital libraries (mentioned above in Section 5.1.1. - 5.1.3). Time spent on Xplor tactics turned out negatively associated with success rate. This implies browsing might not be an effective approach in conducting a specific information search task.

In Task 3, more frequent application of Xplor and EvalI tactics resulted in higher aspectual recall. This suggests that browsing and individual item evaluation would be important in collecting different aspects of information on a particular topic. System design implications related to iterative browsing strategies mentioned above (Section 5.1.1. - 5.1.3) are necessary to support users in exploratory searches.

Additionally, based on SEM, this study comprehensively examined the relationships between four process-related variables (search tactic selections, system support, difficulty, satisfaction with search tactic application) and two output measures (aspectual recall and satisfaction with search results). User perceptions of system support, difficulty and

satisfaction with search tactics are closely associated with each other. To enhance users' experience with a search process, all three concepts of support, difficult, and satisfaction level should be considered in designing digital library systems. This study empirically proved that process variables would directly or indirectly influence users' satisfaction with search outputs. Also, the SEM analysis confirms the effects of search tactic selection and system support on aspectual recall in Task 3.

5.2. Methodological Implications in interactive IR

This dissertation is one of the exploratory studies that comprehensively investigated user engagement and system support at the search tactic level focusing on search process.

At the granularity of search tactics, this study empirically 1) analyzed users' search tactic selections, 2) traced search tactic occurrence probabilities in a session, 3) estimated search tactic transitions, and 4) identified different groups of sessions based on search tactic application characteristics. Multiple methods were employed to comprehensively look into user engagement in a search process, including descriptive statistics, kernel regression, Markov switching, hierarchical clustering, and MDS mapping. From this analysis, this study uses a range of variables that can be used for interactive IR research as follows:

- Frequency of search tactics
- Time spent on applying search tactics

- Search tactic occurrence probability in a single session (estimated from nonparametric regression)
- Transition probability between search tactics
- Dissimilarity between sessions calculated based on search tactic application

These behavioral variables could be practical, useful in interactive IR research.

Previously, interactive IR research has relied widely on usability variables, such as page visits (views), session time, clicks, and page dwell time, to represent user interactions with the system. Search tactic based measures suggested in this study include various interactions comprehensively, ranged from query creation and reformulation, search result evaluation, browsing, accessing forward and backward, and to individual item evaluation. More importantly, the benefit of search tactic based research is the involvement of user intention whereas it is sometimes hard to interpret underlying users' intention from previous behavioral variables. Moreover, the variables suggested herein involve more information about search process, such as transition and probability change. These variables can be incorporated in interactive IR evaluation to better represent user engagement in a search process at the micro-level.

At the users' perceptual level, this study measured system support for each type of search tactic. This study also attempted to measure the degree of difficulty in search tactic application. Additionally, the author tried to measure users' satisfaction level for search tactic application. All these efforts were made to assess a search process, which comprises with sequences of different types of search tactics. Measuring user perceptions of search tactic application could be a compelling method to evaluate the quality of a

search process. For example, more system support and less difficulty in applying search tactics could be interpreted as better search process. Interactive IR research has used users' ratings to collect various data in search tasks, such as task difficulty, search result quality, perceived usability, attention, or preference. These ratings are usually used to assess an entire search task rather than sub-tasks of a session. The uniqueness of this study lies in that it attempted to measure search process by measuring user perceptions of sub-tasks, which are equivalent to search tactics in this study. The variables of perceived system support and perceived difficulty for search tactics can be used as a way to assess the quality of search process. Also, user satisfaction was measured for application of different types of search tactics at the affective level. Of course, it is a controversial issue whether user survey on measuring satisfaction level would be valid or reliable. However, survey is still one of feasible, easy ways to measure users' subjective feelings on search process. This study also showed some positive relationship between search output and satisfaction level. In this way, the author tried to provide new approaches to measure user-system interactions as well as the quality of search process in the context of digital libraries.

5.3. Limitations of the Study

This study has several limitations in its research design, data collection, and data analysis. First, sample size was insufficient to explore various aspects of user engagement and to generalize the findings. Sixty participants do not represent the entire user group of digital libraries, even though the study analyzed 5,465 tactics observed in 180 search sessions.

Also, the sample of this study includes only university students. In reality, most digital libraries are open to the public, but student users are a small portion of the entire digital library users. Therefore, the results of the study cannot be extended to understand general public users' search behavior.

Second, system support was not objectively measured. Instead, it was measured by user perception using a Likert scale. The best way to objectively measure system support is to count frequencies of system feature uses, such as counting advanced search uses or query expansion uses. However, frequencies of system feature uses are not measured in this study. Because of the limited data collection resources, it was not possible to objectively count frequencies of all types of system feature uses. Since transaction log data do not tell precisely different types of system feature uses in detail, the author could not include frequency of system feature uses as a variable to indicate system support. Both an eye-tracker and more structured think-aloud instruction are imperative to objectively count users' actual uses of system features. The only option that this study could select was a Likert scale to measure system support level. Physiological sensory measures were not included, such as galvanic skin response, electromyogram, and electroencephalography, due to the limited research resources.

Third, reasons underlying users' search tactic selections were not analyzed sufficiently. This study described how users select search tactics and what search tactic patterns look like from the quantitative analysis of search tactics. As the focus of the study is on quantitative modeling of search tactic patterns, qualitative analysis of the think-aloud protocols is not included yet in this dissertation. The author plans on further research that

qualitatively identifies user intention of search tactic selection by analyzing think-aloud protocols.

Fourth, this study is based on the analysis of data generated from single search sessions. However, in real information seeking situations, multiple search sessions are also very common. As the data collection was limited to a single visit of each subject, multiple search sessions were not investigated. In addition, each task imposed a time restriction, 5 or 8 minutes for each task. Five minutes seemed to be appropriate for Task 1 and 2, but eight minutes might be insufficient for Task 3.

Fifth, contextual factors were not analyzed in this study, except search task type. The author collected data of several contextual variables, such as different aspects of user knowledge (e.g., search skills, topic knowledge, system familiarity, etc.) and user characteristics (e.g., demographic information, self-efficacy, etc.). The effects of contextual factors are not the interest of this study. This study is designed as an exploratory study, rather than examining factors affecting search tactic patterns. The author plans on a next study examining relationships between various factors and search tactic patterns.

5.4. Chapter Summary

This chapter discusses unique search tactic patterns, implications for the design of system, and implications for methodology. First, the author reviewed users' unique search behavior in digital libraries, including search tactic selection, search tactic probability

changes in a session, and frequent and infrequent search tactics. For each section of the discussion, some implications for digital library system design were suggested. In addition, the limitations of the study were addressed in terms of generalizability, limited variables, lack of qualitative analysis, multiple session analysis, and exclusion of contextual variables.

CHAPTER 6. CONCLUSIONS

This dissertation investigated users' search tactic application and system support in the context of digital libraries. This study assumed that a user engages in an IR process by applying different types of search tactics, and the system supports users' search tactic applications in certain ways. Also, in an attempt to assess the quality of a search process, this study measured users' perceived system support for and difficulty in applying search tactics. Moreover, this study examined how user search tactic application and system support would influence search outputs in different search task situations.

This study empirically answered the research questions based on the analysis of search sessions from sixty subjects.

RQ 1. How do users engage in a search process by applying different types of search tactics while conducting different search tasks?

Users' search tactic patterns were analyzed to answer RQ 1. Users showed different patterns of search tactics by task type. Frequencies of AccF, EvalR, and Xplor tactics were relatively high in known-item search tasks while frequencies of AccB and EvalR tactics were high in specific information search tasks. In particular, iterative search result evaluation strategies were frequently used in specific information search tasks. In exploratory searches, users spent the most time on evaluating individual items. Search tactic selection probability showed different patterns within a session. In the starting phase, Creat and Xplor tactics were most frequently used as a way to initiate a search session. In the beginning of a session, Creat tactics were more frequently selected in Task

1 and 2, while Xplor tactics were preferred in Task 3. In the middle phase, EvalR tactics were dominantly selected in Task 2 while EvalI tactics were most frequently used in Task 3. In the ending phase, high probability of EvalI was observed in Task 1. In Task 2 and 3, the ending phase showed similar search tactic selection patterns with the middle phase. Transition analysis identified frequently applied paths of search tactics. In Task 1, both searching and browsing were frequently used, but few iterations were observed. In Task 2, iterative search result evaluation was dominantly selected. In Task 3, two iterative patterns of search strategies were frequently applied – iterative browsing and iterative result evaluation. In addition, selection of search tactics was used to group sessions with similar patterns.

RQ 2. How does the system support users to apply different types of search tactics?

This study identified types of system supports provided by LOC-DL. For each type of search tactic, LOC-DL supported users' application of search tactics by providing different system features. Also, the study tried to evaluate the quality of search process by measuring system support, difficulty, and satisfaction with regard to application of search tactics. It turned out that users experienced relatively sufficient system support for the application of AccF, AccB, and Xplor tactics, but less support for Mod and EvalI tactics. Quite the reverse, users rated high level of difficulty in applying Mod and EvalI tactics, and perceived relatively easy to apply AccF, AccB, and Xplor tactics. As to satisfaction with search tactic application, users were more satisfied with applying AccF and AccB tactics, whereas less satisfied with application of EvalI tactics. Positive correlation was

observed between system support and satisfaction, but negative correlation between system support and difficulty.

RQ 3. How do users' search tactic application and system support for different types of search tactics affect search outputs?

In Task 1, the application of Creat and AccF tactics positively influenced session efficiency, whereas AccB tactics did negatively. Searching is more efficient than browsing to complete a known-item search task quickly. In Task 2, the frequency of EvalI positively affected success rate, whereas the frequency of Xplor did negatively. Time spent on EvalR positively affected success rate, while time spent on Xplor or Mod did negatively. This suggests iterative search result or item evaluation would be more effective than browsing in achieving specific information search tasks. In Task 3, the frequency of EvalI and Xplor tactics positively affected aspectual recall, and the frequency of EvalI also positively affected satisfaction with search results. Similarly, time spent on EvalI and Xplor tactics showed a significant positive influence on aspectual recall. Perceived system support for Xplor tactics positively affected aspectual recall. SEM results delineate comprehensive relationships between search tactic application, system support, difficulty for search process, and search outputs. Search tactic selection and system support positively affected aspectual recall. Also, system support, satisfaction with search tactics, and aspectual recall influenced satisfaction with search results.

The overall goal of this dissertation is to understand interactions between users and digital library systems focusing on search tactic application during the search process. This research explored search sessions in digital libraries at the granularity of search tactic. This study also has generated several system design implications for digital library systems as well as methodological implications.

As mentioned in Chapter 5, this study has several limitations, and further research is needed to better understand user-system interactions in different search task situations. In the future, the author will expand this study to more participants with a variety of tasks in real settings to better generalize the results. Foremost, contextual factors will be examined in relation to users' search tactic selections and patterns. In the data collection stage, the author measured several variables about user characteristics and user knowledge. These contextual variables will be incorporated in the future analysis to investigate how these factors influence users' search tactic application. The next study will investigate more dynamic relationships among contextual factors, task types, search tactic patterns, and search outputs. More importantly, user attitude and knowledge will be included in the future study including self-efficacy, search skills, and domain knowledge. In this way, future studies will cover a variety of factors more extensively to portray the entire picture of user engagement and system support in digital libraries. Additionally, the author has a plan to use an eye-tracker to objectively measure user-system interactions.

In conclusion, this study has contributed to the understanding of IR sessions at the micro level in using digital library systems. The study tried to quantitatively model search tactic patterns in different task types, measured system support, difficulty, and satisfaction, and

investigated relationships between search process and search output. The findings of this study yield several insights into the design of IR systems and suggest methodological implications in interactive IR research.

REFERENCES

- Agichtein, E., Brill, E., & Dumais, S. (2006). Improving Web Search Ranking by Incorporating User Behavior, in Proceedings of the ACM Conference on Research and Development on Information Retrieval (SIGIR 2006), pp. 19–26.
- Ahn, J. & Brusilovsky, P. (2009). Adaptive visualization of search results: Bringing user models to visual analytics. *Information Visualization*, 8(3):167-179.
- Al-Maskari, A. & Sanderson, M. (2010). A review of factors influencing user satisfaction in information retrieval. *Journal of American Society for Information Science & Technology*, 61(5):859–868.
- Arapakis, I. & Jose, J. (2008) Affective Feedback: An investigation of the role of emotions during an information seeking process. In SIGIR 2008. Proceedings of the 31st Annual ACM SIGIR International Conference on Research and Development in Information Retrieval. New York: ACM.
- Arguello, J., Wu, W.C., Kelly, D., & Edwards, A. (2012). Task complexity, vertical display and user interaction in aggregated search. In Proceedings of SIGIR 2012, Portland, OR.
- Aslam, J.A., Pavlu, V. & Yilmaz, E. (2006). A statistical method for system evaluation using incomplete judgments. In Proceedings of SIGIR 2006, pp. 541-548.
- Baroudi, J.J., Olson, M.H., & Ives, B. (1986). An empirical study of the impact of user involvement on system usage and information satisfaction. *Communications of the ACM*, 29(3), 232-238.

- Bates, M.J. (1979). Information search tactics. *Journal of the American Society for Information Science*, 30(4), 205–214.
- Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13, 407-424.
- Bates, M.J. (1990). Where should the person stop and the information search interface start? *Information Processing and Management*, 26, 575-591.
- Bates, M.J. (2002). Toward an integrated model of information seeking and searching: studies of information seeking in context. *New Review of Information Behavior Research*, 3: 1-16.
- Battleson, B., Booth, A., & Weintrop, J. (2001). Usability testing of an Academic Library Web Site: A Case Study. *Journal of Academic Librarianship*, 27(3): 188-198.
- Beaulieu, M., Robertson, S., and Rasmussen, E. (1996). Evaluating Interactive Systems in TREC. *Journal of the American Society for Information Science*, 47(1): 85-94.
- Belkin, N. J. (1993). Interaction with texts: Information retrieval as information seeking behavior. In G. Knorz, J. Krause, & C. Womser-Hacker (Eds.), *Information retrieval '93: Von der Modellierung zur Anwendung* (pp. 55-66). Konstanz: Universitaetsverlag Konstanz.
- Belkin, N.J. (1996). Intelligent information retrieval: whose intelligence? In J. Krause, M. Herfurth, and J. Marx (Eds). *Herausforderungen an die Informationswirtschaft. Informationsverdichtung, Informationsbewertung und Datenvisualisierung*,

- Proceedings of the 5th International Symposium for Information Science (ISI '96), (pp. 25-31). Konstanz: Universitätsverlag Konstanz.
- Belkin, N. (2010). On the evaluation of interactive information retrieval systems. In *The Janus Faced Scholar, a Festschrift in Honour of Peter Ingwersen*. Accessed from: www.issi-society.info/peteringwersen/pif_online.pdf
- Belkin, N. J. & Vickery, A. (1985) *Interaction in information systems: a review of research from document retrieval to knowledge-based systems*, London, British Library.
- Belkin, N.J., Bierig, R., & Cole, M. (2008). Is Relevance the Right Criterion for Evaluating Interactive Information Retrieval. In: *Proceedings of the ACM SIGIR 2008 Workshop on Beyond Binary Relevance: Preferences, Diversity, and Set-Level Judgments*. <http://research.microsoft.com/~pauben/bbr-workshop>
- Belkin, N.J., Cole, M., & Liu, J. (2009). A Model for Evaluation of Interactive Information Retrieval. *SIGIR Workshop on the Future of IR Evaluation*, July 23, 2009, Boston.
- Belkin, N. J., Cool, C., Stein, A., & Thiel, U. (1995). Cases, scripts and information-seeking strategies: On the design of interactive information retrieval systems. *Expert Systems with Applications*, 9, 379–395.
- Belkin, N.J., Marchetti, P.G., & Cool, C. (1993). BRAQUE: Design of an interface to support user interaction in information retrieval. *Information Processing and Management*, 29(3): 325-344.

- Blandford, A., Keith, S., Connell, I., & Edwards H. (2004). Analytical Usability Evaluation for Digital Libraries: A case study. Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries, pp. 27-36.
- Borgman, C.L., Gilliland-Swetland, A.J., Leazer, G.H., Mayer, R., Gwynn. D., Gazan, R., & Mautone, P. (2000). Evaluating digital libraries for teaching and learning in undergraduate education: a case study of the Alexandria Digital Earth ProtoType (ADEPT). *Library Trends*, 49 (2): 228-250.
- Borlund, P. (2000). Experimental Components for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation*, 56(1): 71-90.
- Borlund, P. (2003). The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3). Available from: <http://informationr.net/ir/8-3/paper152.html>
- Borlund, P. and Ruthven, I. (2008). Introduction to the special issues on evaluating interactive information retrieval systems. *Information Processing and Management*, 44: 1-3.
- Brink, T., Gergle, D., & Wood, S. D. (2002). *Designing Web sites that work: Usability for the Web*. San Francisco: Morgan Kaufmann.
- Brophy, P. (2006). *Measuring Library Performance: Principles and Techniques*. Facet Publishing: London, UK.
- Brophy, P. & Wynne, P.M. (1997). *Management Information Systems and Performance measurement for the electronic library (MIEL2)*. Final report [online]. URL: <http://www.ukoln.ac.uk/services/elib/papers/supporting/pdf/mis.pdf>

- Brusilovsky, P., Chavan, G., & Farzan, R. (2004) Social adaptive navigation support for open corpus electronic textbooks. In: De Bra, P., Nejdl, W. (eds.) Proc. of Third International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2004). Lecture Notes in Computer Science, 3137: 24-33
- Buttenfield, B. (1999). Usability evaluation of digital libraries. *Science & Technology Libraries*, 17 (3-4), 39-59.
- Byström, K. (2002). Information and information sources in tasks of varying complexity. *Journal of the American Society for Information Science and Technology*, 53(7), 581- 591.
- Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management*, 31, 191-213.
- Callan, J., Allan, J., Clarke, J. L.A., Dumais, S., Evans, D.A., Sanderson, M., & Zhai, C. (2007). Meeting of the MINDS: An information retrieval research agenda, *SIGIR Forum*, 44, pp. 25–34.
- Campbell, R.J. (2009). *Campbell's psychiatric dictionary*. Ninth Edition. Oxford University Press, New York.
- Candela, L., Castelli, D., Pagano, P., Thanos, C., Ioannidis, Y., Koutrika, G., Ross, S., Schek, H., & Schuldt, H. (2007). Setting the Foundations of Digital Libraries: The DELOS Manifesto. *D-Lib Magazine*, 13(3/4). Available from: <http://www.dlib.org/dlib/march07/castelli/03castelli.html>
- Catledge, L. D., & Pitkow, J. E. (1995). Characterizing browsing strategies in the World-Wide Web. *Computer Networks and ISDN systems*, 27(6), 1065-1073.

- Chen, H. & Dumais, S. (2000). Bringing order to the Web: Automatically categorizing search results. In Proceedings of 2000 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 145-152.
- Chen, H.-M., & Cooper, M.D. (2002). Stochastic modeling of usage patterns in a web based information system. *Journal of the American Society for Information Science and Technology*, 53(7): 536–548.
- Cheng, J., Hu, X., & Heidorn, P.B. (2010). New Measures for the Evaluation of Interactive Information Retrieval Systems: Normalized Task Completion Time and Normalized User Effectiveness. ASIST 2010, October 22-27, Pittsburgh, PA, USA.
- Chowdhury, A., & Soboroff, I. (2002). Automatic evaluation of World Wide Web search services, In Proceedings of SIGIR 2002, pp. 421–422.
- Cleverdon, C. (1967). The Cranfield tests on index language devices. *Aslib Proceedings*, 19: 173–192.
- Cleverdon, C. (1991). The Significance of the Cranfield Tests on Index Languages. Paper presented at the 14th annual international ACM SIGIR conference on Research and development in information retrieval.
- Cleverdon, C., Mills, J., & Keen, E. (1966). Factors determining the performance of indexing systems. Cranfield, UK: Aslib Cranfield Research Project.
- Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences* (2nd Edition), Lawrence Earlbaum Associates, Hillsdale, NJ.

- Cole, M., Liu, J., Belkin, N.J., Bierig, R., Gwizdka, J., Liu, C., Zhang, J., & Zhang, X. (2009). Usefulness as the Criterion for Evaluation of Interactive Information Retrieval. In: Proceedings of the third Workshop on Human-Computer Interaction and Information Retrieval. <http://cuaslis.org/hcir2009>.
- Cool, C. & Belkin, N. J. (2002). A classification of interactions with information. In Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (pp. 1-15). Greenwood Village, CO: Libraries Unlimited.
- Cool, C. & Xie, I. (2010). Affective utterances as contextual feedback in interactive information retrieval: examples from help-seeking interactions. In Proceedings of the third symposium on Information interaction in context.
- Cutrell, E. and Guan, Z. What are you looking for? an eye-tracking study of information usage in Web search. ACM Conference on Human Factors in Computing Systems (CHI 2007).
- Dillon, A. (1999). Evaluating on TIME: a framework for the expert evaluation of digital library interface usability. *International Journal on Digital Libraries*, 2(2/3), 170-7.
- Diriye, A., Blandford, A., & Tombros, A. (2010). When is system support effective? IliX '10 Proceedings of the third symposium on Information interaction in context, Pages 55-64. ACM, New York

- Du, J. T. & Spink, A. (2011). Toward a web search model: integrating multitasking, cognitive coordination, and cognitive shifts. *Journal of the American Society for Information Science and Technology*, 62(8), 1446-1472.
- Dumais, S.T. & Belkin, N.J. (2005). The TREC interactive tracks: Putting the user into search. In E.M. Voorhees & D.K. Harman (Eds.), *TREC: Experiment and evaluation in information retrieval* (pp. 123-152). Cambridge, MA: The MIT Press.
- Eliassen, K., McKinstry, J., & Fraser, B.M. (1997). Navigating Online Menus: A Quantitative Experiment. *College & Research Libraries*, 58: 509-516.
- Ellis, D. (1989). A behavioural approach to information retrieval design. *Journal of Documentation*, 24(4): 373-389.
- Ellis, D., Cox, D. & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, 49: 356-369.
- Ellis, D. & Haugan, M. (1997). Modeling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, 53(4): 384-403.
- Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. Second edition. New York: Marcel Dekker.
- Fox, E.A., & Urs, S.R. (2002). Digital libraries. In Blaise, C. ed. *Annual Review of Information Science and Technology*. Medford, NJ: Information Today, 36: 503-589.

- Fox, E.A., Neves, F.D., Yu, X., Shen, R., Kim, S. & Fan, W. Exploring the computing literature with visualization and stepping stones & pathways. *CACM* 49(4): 52-58, 2006.
- Fuhr, N., Hansen, P., Mabe, M., Micsik, A., Sølvsberg, I. (2001). Digital libraries: a generic classification and evaluation scheme. *Lecture Notes in Computer Science*, 2163, 187-199.
- Fuhr, N., Tsakonas, G., Aalberg, T., Agosti, M., Hansen, P., Kapidakis, S., Klas, C., Kovács, L., Landoni, M., Micsik, A., Papatheodorou, C., Peters, C., & Sølvsberg, I. (2007). Evaluation of digital libraries. *International Journal on Digital Libraries*, 8, 21-38.
- Fujikawa, K., Joho, H., & Nakayama, S. (2013) Tempo of search actions to modeling successful sessions. In: *Proceedings of the 34th European Conference on Information Retrieval (ECIR 2013)*, pp. 718-721, Moscow.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: strategies for qualitative research*, Chicago: Aldine Transaction.
- Golovchinsky, G., Dunnigan, A., & Diriye, A. (2012, May). Designing a tool for exploratory information seeking. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 1799-1804). ACM.
- Grove, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd Ed.). New Jersey: John Wiley & Sons, Inc.

- Guo, Q., Murphy, A., Deemer, S., & Agichtein, E. (2008). Modeling user interactions for automatic library search evaluation: Preliminary results. In *Proceedings of the JCDL* (Vol. 8).
- Gutwin, C., Paynter, G., Witten, I., Nevill-Manning, C., & Frank, E. (1999). Improving browsing in digital libraries with keyphrase indexes. *Decision Support Systems*, 27: 81-104.
- Gwizdka, J. (2010). Distribution of cognitive load in Web search. *Journal of the American Society for Information Science* 61(11): 2167–2187.
- Hammill, S. (2003). Usability testing at Florida International University Libraries: what we learned. *Electronic Journal of Academic and Special Librarianship*, 4(1).
- Han, S., Yue, Z., & Daqing H. (2013). Automatic detection of search tactic in individual information seeking: A Hidden Markov Model approach. *iConference 2013 Proceedings* (pp. 712-716). doi:10.9776/13330
- Harman, D. K. & Voorhees, E. M. (2006): TREC: An Overview. In: *Annual Review of Information Science and Technology* 40, pp. 113-156
- Harper, S., Michailidou, E., & Stevens, R. (2009). Toward a definition of visual complexity as an implicit measure of cognitive load. *ACM Transactions on Applied Perception*, 6(2), 1–18.
- Hearst, M.A. (2006) Clustering versus faceted categories for information exploration. *Communication of the ACM* April Vol 49, No.4
- Hendry, D. G., & Harper, D. J. (1997). An informal information-seeking environment. *JASIS*, 48(11), 1036-1048.

- Hersh, W. (2009). *Information retrieval: A health and biomedical perspective*. New York: Springer.
- Hersh, W. & Over, P. (2000). SIGIR workshop on interactive retrieval at TREC and beyond. *SIGIR Forum*, 34(1), 24-27.
- Hert, C. (2001). User-centered evaluation and its connection to design. In McClure, C.R., & Bertot, J.C. (Eds.), *Evaluating networked information services: Techniques, policy, and issues* (pp.155–173). Medford: Information Today.
- Hildreth, C. R. (2001) Accounting for users' inflated assessments of on-line catalogue search performance and usefulness: an experimental study. *Information Research*, 6(2), <http://informationr.net/ir/6-2/paper101.html>
- Hill, L.L., Dolin, R., Frew, J., Kemp, R.B., Larsgaard, M., Montello, D.R., Rae, M.-A., & Simpson, J. (1997). User evaluation: Summary of the methodologies and results for the Alexandria Digital Library, University of California at Santa Barbara. *Proceedings of 60th ASIST Annual Meeting*. (pp. 225–243, 369). Medford, NJ: Information Today
- Hix, D., & Hartson, R. (1993). *Developing user interfaces: Ensuring usability through product and process*. New York: John Wiley.
- Hoyle, R. & Panter, A. (1995). Writing about structural equation models. In: Hoyle R (Ed.) *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks: Sage, pp. 100-119.
- Hsieh-Yee, I. (1998). Search tactics of Web users in searching for texts, graphics, known items and subjects: A search simulation study. *Reference Librarian*, 60, 61-85.

- Huang, C., & Xie, I. (2011). Help feature interactions in digital libraries: Influence of learning styles. In Proceedings of ASIST 2011 Annual Meeting, New Orleans, LA, October 9-12, 2011. Online.
- Hung, T.-Y. (2005). Search moves and tactics for image retrieval in the field of journalism: A pilot study. *Journal of Educational Media & Library Science*, 42(3), 329–346.
- Ingwersen, P. (1992). *Information Retrieval Interaction*. London: Taylor Graham.
- Ingwersen, P. & Järvelin, K. (2005). *The turn: integration of information seeking and retrieval in context*. Dordrecht: Springer.
- International Standard Organization (ISO). (1998) 'ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)', Part 11: Guidance on Usability Specification and Measures. Technical report.
- Järvelin, K. (2009). Explaining User Performance in Information Retrieval: Challenges to IR evaluation. In: L. Azzopardi et al. (Eds.), *Proceedings of the 2nd International Conference on the Theory of Information Retrieval*. 2009, Heidelberg: Springer, *Lecture Notes in Computer Science*, vol. 5766, pp. 289-296.
- Jeng, J. (2005). Usability Assessment of Academic Digital Libraries: Effectiveness, Efficiency, Satisfaction, and Learnability. *Libri*, 55: 96–121.
- Jeng, J. (2006). *Usability of the Digital Library: An Evaluation Model*. Unpublished dissertation, Rutgers University.

- Joho, H. (2009) Cognitive Effects in Information Seeking and Retrieval. In: Proceedings of the Workshop on Contextual Information Access, Seeking and Retrieval Evaluation, ECIR'09, Toulouse, France.
- Joo, S. & Lee, J. (2011). Measuring the Usability of Academic Digital Libraries: Instrument Development and Validation. *The Electronic Library*. 29(4): 523-537.
- Joo, S., Lin, S., & Lu, K. (2011). A Usability Evaluation Model for Academic Library Websites: Efficiency, Effectiveness and Learnability. *Journal of Library and Information Studies*, 9(2), 11-26.
- Joo, S. & Xie, I. (2013). Evaluation Constructs and Criteria for Digital Libraries: A Document Analysis. In Cool, C. & Ng, K.B. (Eds). *Recent Developments in the Design, Construction and Evaluation of Digital Libraries*. Hershey, PA: IGI Global. (Pages 126-140).
- Käki, M. (2004). Proportional search interface usability measures. *Proceedings of NordiCHI 2004*.
- Käki, M. (2005). Findex: Search result categories help users when document ranking fails. *Proceedings of the Conference on Human Factors in Computing Systems*, 131–140
- Kang, H. & Schneiderman, B. (2000). Visualization Methods for Personal Photo Collections: Browsing and Searching in the PhotoFinder, *ACM CIKM 2000*, New York, USA, pp. 1539- 1542.

- Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1-2), 1-224.
- Kengeri, R., Seals, C.D., Reddy, H.P., Harley, H.D., & Fox, E.A. (1999). Usability study of digital libraries: ACM, IEEE-CS, NCSTRL, and NDLTD. *International Journal on Digital Libraries*, 2(2/3), 157-169.
- Kenney, A.R., Sharpe, L.H., & Berger, B. (1998). Illustrated book study: Digital conversion requirements of printed illustration. In C. Nikolaou & C. Stephanidis (Eds.), *Research and advanced technology for digital libraries: Proceedings of the Second European Conference* (pp. 279-293). Berlin, Germany: Springer.
- Kent, R.E & Bowman, C.W. (1995). *Digital Libraries, Conceptual Knowledge Systems and the Nebula Interface*. Technical Report, University of Arkanaas.
- Kim, K-S., & Allen, B. (2002). Cognitive and Task Influences on Web Searching Behavior. *Journal of the American Society for Information and Science*, 53(3): 109-119.
- Kim, Y.H., & Kim, H.H. (2008). Development and validation of evaluation indicators for a consortium of institutional repositories: A case study of dcollection. *Journal of the American Society for Information Science and Technology*, 59(8): 1282-1294.
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42, 361–371.

- Kuhlthau, C. C. (1993). *Seeking meaning: A process approach to library and information services*. New Jersey: Ablex Publishing.
- Kules, B., Capra, R., Banta, M. and Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, Austin, TX, USA, June 15-19, 2009: JCDL '09 (pp. 313-322). New York, NY: ACM Press.
- Kules, B. & Capra, R. (2011). Visualizing Stages During an Exploratory Search Session. In *Proceedings of HCIR 2011*, October 20, 2011, Mountain View, CA, USA.
- Kyrillidou, M., & Giersch, S. (2005). Developing the DigiQUAL protocol for digital library evaluation. In M. Marlino, T. Summer, & F. Shipman (Eds.), *Proceedings of the Fifth ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 172-173). New York: ACM Press.
- Larsen, L. (2002). The DLib Test Suite and Metrics Working Group: Harvesting the Experience from the Digital Library Initiative. D-Lib Working Group on Digital Library Metrics Website. Available from:
http://www.dlib.org/metrics/public/papers/The_Dlib_Test_Suite_and_Metrics.pdf
- Lee, J., Paik, W. & Joo, S. (2012). Information resource selection of undergraduate students in academic search tasks" *Information Research*, 17(1) paper511. [Available at <http://InformationR.net/ir/17-1/paper511.html>]

- Li, Y. (2010). An Exploration of the Relationships between Work Tasks and Users' Interaction Performance. Proceedings of the 73rd ASIS&T Annual Meeting, Oct. 2010, Pittsburgh, PA, USA.
- Li, Y., & Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management*, 44(6), 1822-1837.
- Lin, S.J. (2002). Design space of Personalized Indexing: Enhancing successive Web searches for transmuting information problems,” Proceedings of 2002 American Conference on Information Systems. Dallas, TX. August, 9-11, 2002
- Lin, S.J. (2005). Internetworking of factors affecting successive searches over multiple episodes. *Journal of the American Society for Information Science and Technology*, 56(4): 416-436.
- Lin, S.J. & Belkin, N. (2004). Validation of a model of information seeking over multiple search sessions. *Journal of the American Society for Information Science and Technology*, 56(4): 393-415.
- Lin, S. & Xie, I. (2013). Behavioral changes in Transmuting Multi-session Successive searches over the Web. *Journal of American Society for Information Science and Technology*, 64(6): 1259-1283.
- Liu, C., Gwizdka, J., Liu, J., Xu, T., & Belkin, N. J. (2010). Analysis and evaluation of query reformulations in different task types. Proceedings of the 73rd ASIS&T annual meeting, Oct. 2010. Pittsburgh, PA, USA.

- Lopatovska, I. & Mokros, H.B. (2008). Willingness to pay and experienced utility as measures of affective value of information objects: Users' accounts. *Information Processing and Management*, 44: 92-104.
- Lu, J. & Callan, J. (2005). Federated search of text-based digital libraries in hierarchical peer-to-peer networks. *Proceedings of the 27th European Conference on Information Retrieval*, 52-66.
- Marchionini, G. (1995). *Information seeking in electric environment*. New York: Cambridge University Press.
- Marchionini, G. (2006) *Exploratory search: from finding to understanding*. *Communications of the ACM* 49(4): 41-46.
- Makri, S., Blandford, A., & Cox, A. L. (2008a). Using Information Behaviors to Evaluate the Functionality and Usability of Electronic Resources: From Ellis's Model to Evaluation. *Journal of American Society of Information Science and Technology*, 59(14), 2244-2267.
- Marki, S., Blandford, A., & Cox, A.L. (2008b). Investigating the information seeking behavior of academic lawyers: From Ellis's model to design. *Information Processing and Management*, 44(2), 613-634.
- Matusiak, K.K. (2010). *Use of digital resources in an academic environment : a qualitative study of students' perceptions, experiences, and digital literacy skills*. Unpublished Dissertation. University of Wisconsin—Milwaukee.
- Mika, P. (2008). Microsearch: an interface for semantic search, in: *Proceedings of the SemSearch, 2008*, pp. 79–88.

- Miller, M., Choi, G., & Chell, L. (2012). Comparison of three digital library interfaces: open library, Google books, and Hathi Trust. Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries, Pages 367-368.
- Missingham, R. (2003). What Makes Libraries Relevant in the 21st Century? Measuring Digital Collections from Three Perspectives. Information Online Conference. <http://www.nla.gov.au/nla/staffpaper/2003/missingham2.html>.
- Morse, E.L. (2002). Evaluation methodologies for information management systems. D-Lib Magazines, 8(9). Available at: <http://www.dlib.org/dlib/september02/morse/09morse.html>
- Mu, X., Ryu, H., & Lu, K. (2011). Supporting effective health and biomedical information retrieval and navigation: A novel facet view interface evaluation. Journal of biomedical informatics, 44(4), 576-586.
- Nielson, J. (1993). Usability Engineering. Academic Press: Cambridge.
- Noh, Y. (2010). A Study on Developing Evaluation Criteria for Electronic Resources in Evaluation Indicators of Libraries. Journal of Academic Librarianship, 36(1): 41-52.
- O'Brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors, Quality and Quantity, 41(5): 673-690.
- O'Day, V.L., & Nardi, B.A. (2003). An ecological perspective of digital libraries. In Bishop, A.P., Van House, A.A., & Butterfield, B.P. (eds.) *Digital Library Use: Social Practice in Design and Evaluation*. Massachusettes, Cambridge: The MIT Press.

- Oddy, R.N. (1977). Information retrieval through man-machine dialogue. *Journal of Documentation*, 33(1): 1-14.
- Ogilvie, P., & Callan, J. (2003, July). Combining document representations for known-item search. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval* (pp. 143-150). ACM.
- Olston, C. & Chi, Ed H. (2003). ScentTrails: Integrating browsing and searching on the web. *ACM Tarnsactions on Computer-Human Interaction*, 10(3), 177-197.
- Olah, J. (2005). Shifts between search stages during task-performance in mediated information seeking interaction. *Proceedings of the American Society for Information Science and Technology*, 42(1).
- Oulanov, A., & Pajarillo, E.F.Y. (2002). CUNY+ Web: Usability study of the Web-based GUI version of the bibliographic database of the City University of New York (CUNY). *The Electronic Library* 20 (6): 481–87.
- Over, P. (2001). The TREC Interactive Track: an annotated bibliography. *Information Processing and Management*, 37: 369-381.
- Palmquist, R.A., & Kim, K-S. (2000). Cognitive style and on-line database search experience as predictors of Web search performance. *Journal of the American Society for Information Science*, 51(6): 558-566.
- Piorkowski, D., Fleming, S., Scaffidi, C., Bogart, C., Burnett, M., John, B., Bellamy, R. & Swart, C. (2012). Reactive information foraging: An empirical investigation

- of theory-based recommender systems for programmers. In Proceedings of SIG-CHI 2012, ACM, pp. 1471–1480
- Rieh, R. & Xie, H. (2006). Analysis of multiple query reformulations of the Web: The interactive information retrieval context. *Information Processing and Management*, 42, 751-768.
- Robins, D. (2000). Interactive information retrieval: Context and basic notions. *Informing Science*, 3(2), 57– 61.
- Rose, D.E. & Levinson, D. Understanding user goals in Web search. In Proc. WWW 2004, (2004), pp. 13-19.
- Rubin, J. & Chisnell, D. (2008). *Handbook of usability testing : how to plan, design, and conduct effective tests*. Indianapolis, IN : Wiley.
- Saracevic, T. (1991). Individual Differences in Organizing, Searching and Retrieving Information. In Proceedings of the ASIS annual meeting (Vol. 28, pp. 82-86).
- Saracevic, T. (1996a). Relevance reconsidered. In P. Ingwersen & N.O. Pors (Eds.), *Information science: Integration in perspectives*. (pp. 201-218). Copenhagen, Denmark: Royal School of Library and Information Science.
- Saracevic, T. (1996b). Modeling interaction in information retrieval (IR): A review and proposal. *Proceedings of the American Society for Information Science*, 33: 3-9.
- Saracevic, T. (1997). The stratified model of information retrieval interaction: Extension and applications. In Proceedings of the ASIS Annual Meeting, (Vol. 34, pp. 313-327).

- Saracevic, T. (2000). Digital Library Evaluation: Toward an Evolution of Concepts. *Library Trends*, 49(3): 350-369.
- Saracevic, T. (2004). Evaluation of digital libraries: An overview. Presented at the DELOS Workshop on the Evaluation of Digital Libraries. Available from: http://www.scils.rutgers.edu/~tefko/DL_evaluation_Delos.pdf
- Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(3), 1915-1933.
- Savage-Knepshield, P. A., & Belkin, N. J. (1999). Interaction in information retrieval: Trends over time. *Journal of the American Society for Information Science*, 50(12), 1067-1082.
- Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: Toward a dynamic, situational definition. *Information Processing & Management*, 26(6), 755-776.
- Scriven, M. (1991). *Evaluation Thesaurus* (4th edition). Newbury Park: Sage Publications.
- Shackel, B. (1991). Usability – Context, framework, definition, design and evaluation. In Shackel, B & Richardson, S. (Eds.), *Human Factors for Informatics Usability*, Cambridge University Press, Cambridge, UK, p.21-37.
- Shen, R., Vemuri, N.S., Fan, W., Torres, R., & Fox, E.A. (2006). Exploring digital libraries: integrating browsing, searching, and visualization. In *JCDL '06*:

- Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries, pages 1–10, New York, NY, USA, 2006. ACM Press.
- Shim, W. and Kantor, P. (1999). Evaluation of digital libraries: A DEA approach. Proceedings of the 62nd ASIST Annual Meeting: 605-615.
- Shiri, A. A., & Revie, C. (2003). The effects of topic complexity and familiarity on cognitive and physical moves in a thesaurus-enhanced search environment. *Journal of Information Science*, 29(6), 517-526.
- Silverman, B.W (1986). *Density Estimation for Statistics and Data Analysis*, ser. Monographs on Statistics and Applied Probability. New York: Chapman & Hall.
- Skov, M. (2009). *The Reinvented Museum: Exploring Information Seeking Behaviour in a Digital Museum Context*, PhD thesis, Research Programme Information Interaction and Information Architecture, Royal School of Library and Information Science, Denmark.
- Smith, A. G. (2012). Internet search tactics. *Online Information Review*, 36(1), 7-20.
- Spink, A. (1997). Study of interactive feedback during mediated information retrieval. *Journal of the American Society for Information Science*, 48(5): 382-394.
- Spink, A. & Greisdorf, H. (2000). Regions and levels: Measuring and mapping users' relevance judgments. *Journal of the American Society for Information Science and Technology*. 52(2): 161–173.
- Strauss, A. & Corbin, J. (1990). *Basics of Qualitative Research: Grounded theory procedures and techniques*. Sage: Newbury Park.

- Su, L.T. (1992). Evaluation measures for interactive information retrieval. *Information Processing & Management*, 28(4): 503–516.
- Su, L.T. (1994). The relevance of recall and precision in user evaluation. *Journal of the American Society for Information Science*, 45(3), 207–217.
- Takeda, H., Farsiu, S., & Milanfar, P. (2007). Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, 16(2), 349–366.
- Tamine-Lechani L, Boughanem M, Daoud M (2010) Evaluation of contextual information retrieval effectiveness: overview of issues and research. *Knowledge Information System*, 24(1): 1-34.
- Tenopir, C., Wang, P., Zhang, Y., Simmons, B. & Pollard, R. (2008). Academic users' interactions with ScienceDirect in search tasks: Affective and cognitive behaviors. *Information Processing and Management*, 44: 105-121.
- Tidwell, J. (2011). *Designing Interfaces*, Second Edition. O'Reilly Media: Sebastopol, CA, USA.
- Tsakonas, G., Kapidakis, S., & Papatheodorou, C. (2004). Evaluation of user interaction in digital libraries. In M. Agosti, N. Fuhr [eds.] *Notes of the DELOS WP7 Workshop on the Evaluation of Digital Libraries*, Padua, Italy.
- Tullis, T. & Albert, B. (2008). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*. Amsterdam: Elsevier/Morgan Kaufmann.
- Turpin, A. & Scholer, F. (2006). "User performance versus precision measures for simple search tasks". *Proceedings of the 29th Annual international ACM SIGIR Conference on Research and Development in information Retrieval*, pp. 11–18.

- Vakkari, P. (2000a). E-cognition and changes of search terms and tactics during task performance: A longitudinal case study. In Proceedings of the RIAO-2000 Conference, (pp. 894-907). Paris, France: C.I.D. Retrieved January 4, 2008, from http://www.info.uta.fi/vakkari/Vakkari_Tactics_RIAO2000.html
- Vakkari, P. (2000b). Relevance and contributory information types of searched documents in task performance. In Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (pp. 2-9). New York: ACM Press.
- Vakkari, P. (2001). A theory of the task-based information retrieval process. *Journal of Documentation*, 57(1), 44-60.
- Vakkari, P. (2003). Task-based information searching. *Annual Review of Information Science & Technology*, 37, 413-464
- Vakkari, P. (2013). Evaluating Interactive Information Retrieval Systems. *Revista PRISMA. COM*, (19).
- Vakkari, P., & Hakala, N. (2000). Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, 56, 540-562.
- Vakkari, P., & Kekäläinen, J. (2011). An evaluation methodology reflecting the aims of search tools. *TPDL 2011 Workshop on User Oriented Evaluation of Digital Library Interfaces*, Berlin, pp. 1-9.
- Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal: A longitudinal case study. 39(3), 445-463.

- Van House, N. A., Butler, M.H., Ogle, V., & Schiff, L. (1996). User-centered iterative design for digital libraries: The Cypress experience. *D-Lib Magazine*, 2. Available from: <http://www.dlib.org/dlib/february96/02vanhouse.html>
- Voorhees, E.M. (2003). Overview of the TREC 2002 question answering track. In *Proceedings of the Eleventh Text REtrieval Conference (TREC 2002)*, p. 57–68.
- Ward, J.H. (1963). Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, 58, 236–244.
- Ward, J.L., & Hiller, S. (2005). Usability testing, interface design, and portals. *Journal of Library Administration*, 43(1): 155-171.
- Wästlund, E., Norlander, T., & Archer, T. (2008). The effect of page layout on mental workload: A dual-task experiment. *Computers in Human Behavior*, 24(3), 1229–1245.
- White, R W., Kules, B., Drucker, S.M., & Schraefel, M.C. (2006). Supporting exploratory search. *Communications of the ACM*, 49(4), 36–39.
- Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, 55(3), 246-258.
- Wison, T.D. (1999). Models in Information Behaviour Research. *Journal of Documentation*, 55(3), 249–270.
- Wilson, T.D. (2000). Human information behavior. *Informing Science*, 3(2), 49-55

- Xie, I. (2003). Supporting ease-of-use and user control: Desired features and structure of Web-based online IR systems. *Information Processing & Management*, 39, 899-922.
- Xie, I. (2006). Evaluation of digital libraries: criteria and problems from users' perspectives. *Library & Information Science Research*, 28(3), 433-452.
- Xie, I. (2008a). *Interactive information retrieval in digital environments*. Hershey, PA : IGI Pub.
- Xie, I. (2008b). Users' Evaluation of Digital Libraries: Their uses, their criteria, and their assessment. *Information Processing & Management*, 44(3), 1346-1373
- Xie, I. (2009). Dimensions of Tasks: Influences on Information Seeking and Retrieving Process. *Journal of Documentation*. 65(3), 339-366.
- Xie, I. & Benoit, E. (2013). Search Result List Evaluation versus Document Evaluation: Similarities and Differences. *Journal of Documentation*, 69(1): 49 - 80.
- Xie, H. & Cool, C. (2000). Ease of use versus user control: An evaluation of web and non-web interfaces of online databases. *Online Information Review*, 24(2): 102-115.
- Xie, I. & Cool, C. (2009). Understanding Help Seeking Within the Context of Searching Digital Libraries. *Journal of the American Society for Information Science and Technology*, 60(3): 477-494.
- Xie, I. & Joo, S. (2009). Selection of Information Sources: Accessibility of and familiarity with sources, and types of tasks. *Proceedings of the 72th American Society of Information Science Annual Meeting*, 46.

- Xie, I. & Joo, S. (2010a). Transitions in Search Tactics during the Web-based Search Process. *Journal of American Society for Information and Science Technology*. 61(11): 2188-2205.
- Xie, I. & Joo, S. (2010b). Tales from the field: Search strategies applied in Web searching. *Future Internet*, 2(1): 259-281.
- Xie, I. & Joo, S. (2010c). Evaluation Constructs and Criteria for Digital Libraries: A Document Analysis. *Proceedings of TEFKO 2010*, Nov. 2010, New Jersey, USA.
- Xie, I. & Joo, S. (2012) Factors Affecting the Selection of Search Tactics during the Web-based Searching Process. *Information Processing and Management*, 48(2):254-270.
- Xie, I., & Wolfram, D. (2002). State digital library usability: contributing organizational factors. *Journal of the American Society for Information Science & Technology*, 53(13): 1085-1097.
- Yasunaga, Y., Morishima, A., Sodeyama, H., Tadaishi, M., & Sugimoto, S. (2013). Gradation: A pay-as-you-go style hybrid query language for structured and text data. In *Proceedings of iConference 2013* (pp. 209-222).
- Yuan, X. & Belkin, N.J. (2007). Supporting multiple information-seeking strategies in a single system framework. *SIGIR '07 Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 247-254.

- Yue, Z., Han, S., & He, D. (2012). A Comparison of Action Transitions in Individual and Collaborative Exploratory Web Search. In *Information Retrieval Technology* (pp. 52-63). Springer Berlin Heidelberg.
- Zhang, J.(2008). *Visualization for Information Retrieval*. Springer.
- Zhang, J. and Marchionini, G. (2005). Evaluation and evolution of a browse and search interface: relation browser. In *Proc. of the 2005 national conference on digital government research*, 179-188.
- Zhang, X., Li, Y., Liu, J., & Zhang, Y. (2008). Effects of interaction design in digital libraries on user interactions, *Journal of Documentation*, 64(3), 438 - 463.
- Zhang, Y. (2010a). *Developoing a holistic model for digital library evaluation*. PhD Dissertation, Graduate School- New Brunswick Rutgers, The State University of New Jersey.
- Zhang, Y. (2010b). Developing a holistic model for digital library evaluation. *Journal of the American Society for Information Science and Technology*, 61(1), 88-110.
- Zhang, Y., Wang, P., Heaton, A., & Winkler, H. (2012). Health information searching behavior in MedlinePlus and the impact of tasks. In *Proceedings of the ACM International Health Informatics (IHI) Symposium 2012*, (pp. 641-650), Miami, FL.

APPENDIX 1. Pre-questionnaire and Post-Questionnaire

Participant No: _____

Filling out this questionnaire indicates that I am at least eighteen old and I am giving my informed consent to be a participant in this study.

Pre-Questionnaire**Age**
☐ 18-21 ☐ 21-29 ☐ 30-39 ☐ 40-49 ☐ 50-59 ☐ >59
Gender
☐ Female ☐ Male
Native Language
☐ English ☐ non-English
Ethnicity
☐ African American ☐ Asian ☐ Caucasian ☐ Hispanic ☐ Native American ☐ Other
Educational Background (Major)
☐ Undergraduate Program: _____ ☐ Graduate Program: _____
How do you rate your information search skill on the Web:

- ☐ Little knowledge or skills (just learning how to search information on the Web, need lots of help)
☐ Beginner (I need some help to search something on the Web)
☐ Intermediate (Fluent with using commercial search engines like Google and Yahoo.)
☐ Advanced (Fluent with using advanced search functions)
☐ Expert (Good at using advanced search functions, use complex Boolean operators, understand back-end information retrieval mechanisms)

Please rate the frequency with which you use the following by circling the appropriate number. (1=never use, 2=rarely use, 3=occasionally use, 4=often use, 5=use daily)

Type of Systems	Never Use				Use Daily
Web pages	1	2	3	4	5
Web search engines (e.g. Google, Yahoo)	1	2	3	4	5
Library of Congress Digital Collections	1	2	3	4	5
Other digital library please specify:	1	2	3	4	5

Type of Systems	Never Use				Use Daily
Online databases (e.g. EBSCO, ProQuest)	1	2	3	4	5
Library catalogue (e.g. Panther Cat)	1	2	3	4	5
UWM Library Digital Collections	1	2	3	4	5
Other digital library please specify:	1	2	3	4	5

Overall perception of help functions

What do you typically do if you encounter any problems in using an information searching system?

1. ☐ try again 2. ☐ try different approach 3. ☐ consult system Help 4. ☐ ask another person
5. ☐ change systems 6. ☐ give up 7. ☐ Other _____ (please specify)

To what extent do you think help functions of an information searching system are important? Please check one from the following.

1. ☐ not at all 2. ☐ a little 3. ☐ some 4. ☐ some more 5. ☐ extremely

Why do you think that help functions are important or not important?

To what extent do you use help functions of a searching system? Please check one from the following.

1. ☐ never use 2. ☐ rarely use 3. ☐ occasionally use 4. ☐ often use 5. ☐ use every time

Why do you use or not use help functions of a searching system? (just put your reasons, you don't need to write complete sentences.)

How do you learn to use a new searching system when you use it for the first time?

1. ☐ trial and error 2. ☐ consult system Help (e.g. FAQ, search tips, etc.) 3. ☐ ask another person
4. ☐ Other _____ (please specify)

Thank you for completing the questionnaire. After finishing this questionnaire, please email to me at huang22@uwm.edu / sjoo@uwm.edu or bring it with you on the experiment day. Thanks again for your participation.

Please read the following statements carefully!!!!

Activity	Statement	Please rate to what extent the system helped your activity during the information search process				
		Not at all helped (1)	Slightly helped (2)	Somewhat helped (3)	Very helped (4)	Completely helped (5)
Create search terms	This system helped me when I created search terms when using a search box.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Modify search terms	This system helped me change my initial search terms to get better search results.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Choose items from the search results	This system helped me select related items from the search results.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Access to the item	This system helped me access to the item that I found from the search results or browsing categories.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Evaluate the information found	This system helped me judge whether the information of individual items was useful or relevant.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Access back to previous pages	This system helped me move back to the pages that I visited.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Browse different information	This system helped me browse through topic categories or item list.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>

Activity	Statement	Please rate your satisfaction with the following activities during the search.				
		Not at all satisfied (1)	Slightly satisfied (2)	Somewhat satisfied (3)	Very satisfied (4)	Completely satisfied (5)
Create search terms	I was satisfied with the process when I created search terms using a search box.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Modify search terms	I was satisfied with the process when I changed my initial search terms to get better search results.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Choose items from the search results	I was satisfied with the process when I selected related items from the search results.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Access to the item	I was satisfied with the process when I accessed to the item that I found from the search results or browsing categories.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Evaluate the information found	I was satisfied with the process when I judged whether the information of individual items was useful or relevant.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Access back to previous pages	I was satisfied with the process when I moved back to the pages that I visited.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Browse different information	I was satisfied with the process when I browsed through different categories or item list.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>

Activity	Items	Please rate how difficult it was to search the information using the LOC digital library				
		Not at all Difficult (1)	Slightly Difficult (2)	Somewhat difficult (3)	Very Difficult (4)	Extremely Difficult (5)
Create search terms	It was difficult for me to create search terms when using a search box.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Modify search terms	It was difficult for me to change my initial search terms to get better search results.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Choose items from the search results	It was difficult for me to select related items from search results.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Access to the information	It was difficult for me to access to the item that I found from search results or item list.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Evaluate the information found	It was difficult for me to judge whether the information I found was useful.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Access back to previous pages	It was difficult for me to move back to the pages that I visited.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>
Browse different information	It was difficult for me to browse through different topic categories or item list.	(1) <input type="checkbox"/>	(2) <input type="checkbox"/>	(3) <input type="checkbox"/>	(4) <input type="checkbox"/>	(5) <input type="checkbox"/>

*Curriculum Vitae***Soohyung Joo**

School of Information Studies
 University of Wisconsin - Milwaukee
 Milwaukee, WI, 53211
 e-mail: sjoo@uwm.edu

EDUCATION

University of Wisconsin – Milwaukee, *WI, USA*

- PhD Candidate, School of Information Studies, Sep. 2008 ~ Dec. 2013

Yonsei University, *Seoul, Korea*

- Master of Library and Information Science (MLIS), Aug. 2008

EXPERIENCE

School of Information Studies, University of Wisconsin-Milwaukee, *WI, USA*

- Adjunct Lecturer, Sep. 2013 ~ Dec. 2013
- INFO SCI 310 – Introduction to System Analysis and Design (Fall 2013)
- INFO SCI 340 – Human Factors in Information Seeking Process (Fall 2013)

School of Information Studies, University of Wisconsin-Milwaukee, *WI, USA*

- Instructor, Jan. 2012 ~ Dec. 2012
- INFO SCI 440 – Information Architecture III (Fall 2012)
- INFO SCI 440 – Information Architecture III (Spring 2012)

School of Information Studies, University of Wisconsin-Milwaukee, *WI, USA*

- Teaching Assistant, Sep. 2010 ~ Dec. 2011
- INFO SCI 440 – Information Architecture III
- INFO SCI 571 – Information Access and Retrieval
- INFO SCI 110 – Introduction to Information Science

School of Information Studies, University of Wisconsin-Milwaukee, *WI, USA*

- Project Assistant, Sep. 2008 ~ Aug. 2011

REFEREED JOURNAL ARTICLES

Lee, J. & Paik, W., & Joo, S. (2012). Resource Selection of Undergraduate Students for Academic Tasks. *Information Research*. 17(1) : Paper 511.

Xie, I. & Joo, S. (2012). Factors Affecting the Selection of Search Tactics: Tasks, Knowledge, Process and Systems. *Information Processing & Management (IPM)*, 48(2): 254-270.

Joo, S., Lin, S., & Lu, K. (2011). A Usability Evaluation Model for Academic Library Websites: Efficiency, Effectiveness and Learnability. *Journal of Library and Information Studies*. 9(2): 11-26.

Lu, K., Joo, S. & Wolfram, D. (2011). An Investigation of Web Information Resource Distribution in the Field of Information Science. *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics*, 15(1): Paper 1

Joo, S. & Lee, J. (2011). Measuring the Usability of Academic Digital Libraries: Instrument Development and Validation. *The Electronic Library*. 29(4): 523-537.

- Xie, I. & Joo, S. (2010). Transitions in Search Tactics during the Web-based Search Process. *Journal of American Society for Information and Science Technology (JASIST)*, 61(11): 2188-2205.
- Xie, I. & Joo, S. (2010). Tales from the field: Search strategies applied in Web searching. *Future Internet*, 2(1): 259-281.
- Lee, J., Han, S., & Joo, S. (2008). The Analysis of the Information Users' Needs and Information Seeking Behavior in the Field of Science and Technology. *Journal of the Korean Society for Information Management*, 25(2), 127-141.
- Lee, J., Min, J., and Joo, S. (2007). A Study to Analyze the Korean National Knowledge Distribution Status and to Generate Suggestions for Developing a Distribution Model. *Journal of the Korean Society for Information Management*, 24(3): 299-319.

BOOK CHAPTER

- Joo, S. & Xie, I. (2013). Evaluation Constructs and Criteria for Digital Libraries: A Document Analysis. In Cool, C. & Ng, K.B. (Eds). *Recent Developments in the Design, Construction and Evaluation of Digital Libraries*. Hershey, PA: IGI Global. (Pages 126-140).

PRESENTATIONS (Conference Paper, Poster, Forum, Workshop)

- Xie, I., Babu, R., Jeong, W., Joo, S., & Fuller, P. (accepted). Blind users searching digital libraries: Types of help-seeking situations at the cognitive level. Poster accepted at the *iConference 2014*, Mar. 4-7, Berlin, Germany.
- Hu, R., Lu, K., & Joo, S. (2013). Effects of Topic Knowledge and Search Skills on Query Reformulation Behavior. (Paper) In *Proceedings of the 76th Annual Meeting of the Association for Information Science and Technology (ASIS&T 2013)*, Nov. 1-6, Montreal, Quebec, Canada.
- DesArmo, J., You, S., & Joo, S. (2013). Measuring Happiness of US Cities by Mining User-generated Text in Flickr.com. (Poster) In *Proceedings of the 76th Annual Meeting of the Association for Information Science and Technology (ASIS&T 2013)*, Nov. 1-6, Montreal, Quebec, Canada.
- Xie, I., Joo, S., & Bennett-Kapusniak, R. (2013). User Engagement and System Support in the Search Process: User dominated, system dominated and balanced tactics. (Poster) In *Proceedings of the 76th Annual Meeting of the Association for Information Science and Technology (ASIS&T 2013)*, Nov. 1-6, Montreal, Quebec, Canada.
- Xie, I., Joo, S., & Matusiak, K. (2013). Multifaceted Digital Library Evaluation Criteria and Measurements. Paper accepted at *the Conference of Korea Biblia Society for Library and Information Science*. Oct. 23-25, Jeju Island, Korea.
- Lin, S. & Joo, S. (2013). Multi-dimensional Evaluation Model for Academic Institutional Repositories. Poster presented at *ARCHIVES 2013*, Aug. 11-17, New Orleans, LA, USA.
- Kipp, M., Joo, S. & Choi, I. (2013). Classification of Online Resources using User Generated Terms. In *Proceedings of IFLA 2013 Classification and Indexing Section*, Aug. 17-23, Singapore.

- Joo, S. & Xie, I. (2013). How Do User Search Tactic Selections Influence Search Outputs in Exploratory Search Tasks? (Poster) In *Proceedings of Joint Conference on Digital Libraries 2013* (JCDL 2013), Jul. 22-26. Indianapolis, IN, USA.
- Choi, I., Joo, S. & Kipp, M. (2013). Web Resource Organization using Consumer Vocabulary in the Domain of Health Information. Poster presented at the *Fourth North American Symposium on Knowledge Organization (NASKO 2013)*, Jun. 13-14, Milwaukee, WI, USA.
- Xie, I., Joo, S. & Matusiak, K. (2013). Digital Library Evaluation Criteria from Different Stakeholders: A document analysis and Delphi-survey. Paper presented at the *5th Qualitative and Quantitative Methods in Libraries International Conference 2013* (QQML 2013), Jun. 4-7, Rome, Italy.
- Joo, S. & Choi, I. (2013). Estimating User Satisfaction in Academic Libraries: An instrument development and validation. Poster presented at *5th Qualitative and Quantitative Methods in Libraries International Conference 2013* (QQML 2013), Jun. 4-7, Rome, Italy.
- Choi, I., Joo, S. & Kipp, M. (2013). Web Resource Clustering Based on Social Tag Information: A Pilot Study. Poster presented at the *5th Qualitative and Quantitative Methods in Libraries International Conference 2013* (QQML 2013), Jun. 4-7, Rome, Italy.
- Joo, S. & Choi, I. (2013). Assessing Multiple Aspects of Academic Library Services: Collections, Staff, Accessibility and Space. Poster presented at WAAL Annual Conference 2013, April 23-26, Elkhart Lake, WI, USA.
- Joo, S. (2013). Exploring Search Tactic Patterns in Using Academic Digital Libraries. Paper Presented at the *1st SOIS Student Research Conference*, April 19, Milwaukee, WI, USA.
- Joo, S., Wolfram, D. & Song, S. (2013). Nonparametric Estimation of Query Term Size and Token Distribution. (Poster) In *Proceedings of iConference 2013*. Feb. 12-15, Fort Worth, Texas, USA.
- Joo, S. (2013). User Engagement and System Support during the Search Process in Digital Libraries. Poster presented at Doctoral Student Poster Competition at *ALISE 2013*. Jan. 22-25, Seattle, WA, USA.
- Joo, S., & Xie, I. (2012). Exploring Search Tactic Patterns in Searching Digital Libraries. In *The Outreach of Digital Libraries: A Globalized Resource Network* (pp. 349-350). Springer Berlin Heidelberg.
- Joo, S. (2012). Exploring Search Tactics during a Search Session in Digital Libraries. In *Proceedings of International Workshop on Global Collaboration of Information Schools (WIS2012)*. Nov. 16. Taipei, Taiwan.
- Joo, S. & Lu, K. (2012). Structural Analysis of Author Vector Space in Information Sciences. (Poster) In *Proceedings of 75th ASIS&T Annual Meeting*. Baltimore, October 26-30, USA.
- Huang, C., Joo, S. & Xie, I. (2012). Effects of learning styles on the selection of search tactics: A preliminary result (Poster). In *Proceedings of 75th ASIS&T Annual Meeting*. Baltimore, October 26-30, USA.
- Oh, K., Joo, S. & Park, Y. (2012). Which Facets are Easy to Use and Useful in Accessing Consumer Health Information on the Web? (Poster) In *Proceedings of iConference 2012*, Toronto, Canada.

- Xie, I., Joo, S., & Kapusniak, R. (2012). Exploring User Engagement and System Support in Applying Search Tactics. Paper presented at *3rd Annual SLIS-SOIS Forum*, Apr. 2012, Madison, WI, USA.
- Joo, S. & Lee, J. (2011). Assessing Effectiveness of Query Reformulations: Analysis of User-Generated Information Retrieval Diaries (Poster). In *Proceedings of 74th ASIS&T Annual Meeting*. Oct. 9-12. New Orleans, USA.
- Joo, S. & Kipp, M.E.I. (2011). Mining Social Tags to Explore the Semantic Web of Health Information Resources. Presented at *74th ASIS&T AM SIG-MET Workshop*. Oct. 12. New Orleans, USA.
- Joo, S. & Choi, Y. (2011). Dimensions and Types of Consumer Vocabulary in Health Information Organisation. Poster presented at *North America Symposium for Knowledge Organization 2011*, Jun. 15-16. Toronto, Canada.
- Kipp, M.E.I. & Joo, S. (2011). Semantic Structure of Consumer Health Information Web Resources Tagged on Delicious.com. Paper presented at *2nd Annual SLIS-SOIS Forum*, Apr. 2011, Milwaukee, WI, USA.
- Joo, S. & Lu, K. (2011). Factors affecting Satisfaction and Re-visit Intention in Three Different Information Resource Environments: Search Engine, Online Database, and OPAC. Work-in-progress Poster presented at *ALISE 2011*, Jan. 4, San Diego, CA, USA.
- Kipp, M.E.I. & Joo, S. (2010). Application of Structural Equation Modelling in Exploring Tag Patterns: A Pilot Study (Poster). In *Proceedings of 73rd ASIS&T Annual Meeting*. Oct. 22-27, Pittsburgh, PA, USA. [Best Poster Award]
- Joo, S. (2010). How Are Usability Elements - Efficiency, Effectiveness, and Satisfaction - Correlated with Each Other in the Context of Digital Libraries? (Poster) In *Proceedings of 73rd ASIS&T Annual Meeting*. Oct. 22-27, Pittsburgh, PA, USA.
- Jeong, W., Davies, K., Olson, H., Xie, H., Lu, K., Joo, S., Kim, T.Y., & Mu, X. (2010). Consumer Health Web Portal Integrated Navigation. Poster at *the 2010 IFLA Conference*, Gothenburg, Sweden. (author not presenter)
- Joo, S. (2010). Two Multivariate Techniques in Informetrics Studies: A Comparison of MDS and SEM. Paper presented at *4th SOIS Student Research Symposium 2010*, Nov. 6., Milwaukee, WI, USA.
- Kipp, M.E.I. & Joo, S. (2010). The Structure of Tag Space in Library and Information Science. Paper presented at *1st Annual SLIS-SOIS Forum 2010*, Apr. 30, Madison, WI, USA.
- Kipp, M.E.I. & Joo, S. (2010). Co-occurrence Analysis of Delicious.com Tag Data Using Social Network Analysis. Poster presented at *1st Annual SLIS-SOIS Forum 2010*, Apr. 30, Madison, WI, USA.
- Joo, S. (2010). A Web Link Structure of the American Library & Information Science Field: A Pilot Study. (Poster) In *Proceedings of iConference 2010*. Feb. 3-6, Urbana-Champaign, IL, USA.
- Joo, S. (2010). Developing a Usability Measurement Instrument in Academic Digital Libraries. (Poster) In *Proceedings of iConference 2010*. Feb. 3-6, Urbana-Champaign, IL, USA.
- Joo, S. (2010). Effects of Ease of Use, Effectiveness, and Use Frequency on User Satisfaction in Academic Library Website Uses. (Poster) In *Proceedings of iConference 2010*. Feb. 3-6, Urbana-Champaign, IL.

- Xie, I., & Joo, S. (2009). Selection of Information Sources: Types of tasks, accessibility and familiarity with sources. In *Proceedings of 72nd ASIS&T Annual Meeting*, Nov. 2009, Vancouver, Canada.
- Ahn, E., & Joo, S. (2009). An Experiment of Automatic News Article Classification using Naïve-bayes Classifier. Presented at *IIA Information Retrieval Seminar*, UW-Milwaukee, Oct. 5. 2009.
- Joo, S. (2008). Application of ISO 9241-11 Standards in Academic Digital Libraries. Presented at *2nd SOIS Student Research Symposium 2008*, Oct. 21. Milwaukee, WI, USA.

AWARDS & SCHOLASHIPS

- SOIS Teaching Award (2012)
 - School of Information Studies, University of Wisconsin-Milwaukee
- Dissertation Fellowship (2012-2013)
 - Graduate School, University of Wisconsin-Milwaukee
- Doctoral Research Award Grant (2011)
 - School of Information Studies, University of Wisconsin-Milwaukee
- ASIS&T Annual Meeting Best Poster Award (2010)
 - The 73rd Annual Meeting of American Society of Information & Science Technology
- Chancellor's Scholarship (2008-2012)
 - University of Wisconsin-Milwaukee
- Dean's Scholarship (2008-2011)
 - School of Information Studies, University of Wisconsin-Milwaukee
- Graduate School Travel Awards (2009)
 - Graduate School, University of Wisconsin-Milwaukee
- National Research Foundation Fellowship (2007-2008)
 - Korea Research Foundation
- Graduate Student Assistant (2006-2007)
 - Department of Library & Information Science, Yonsei University
- Yonsei Graduate School Scholarship (2006)
 - Graduate School, Yonsei University