

THE EFFECTS OF DOMAIN EXPERTISE ON EXPLORATORY INFORMATION  
SEARCH AND TOPIC LEARNING IN SOCIAL SEARCH ENVIRONMENT

BY

RUOGU KANG

THESIS

Submitted in partial fulfillment of the requirements  
for the degree of Master of Science in Human Factors  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Assistant Professor Wai-Tat Fu

## **ABSTRACT**

The goal of this thesis is to expand the traditional information processing model into the social space by investigating the influence of domain expertise on the use of social information systems. A laboratory-based experiment was conducted to examine the information seeking, sharing, and learning processes of domain experts and novices using a traditional search engine and a social tagging system. Empirical data on information behavior, search strategies, information content and knowledge change were recorded and analyzed. Results showed that domain experts collected and shared more information than novices, providing support to the hypothesis that domain experts benefit more from social information systems. Results also showed that the social information system helped domain novices to find general information and facilitated knowledge learning on novices, but the system did not help them to find as much domain-specific knowledge as domain experts, providing support to the hypothesis that domain knowledge is critical for successful utilization of social cues provided by social information systems. Results from the current study also support the notion that there is a dynamic interaction between knowledge-in-the-head and knowledge-in-the-social-web while people are searching in a social information system. Although information seekers are more and more reliant on accessing information from the World Wide Web, the current results suggest that domain expertise is still important for information seekers to successfully find relevant information in both traditional and social information environments. Implications on the design of future social information systems that facilitate exploratory search are discussed.

## **ACKNOWLEDGEMENT**

First of all, I would like to express my sincere gratitude to my adviser, Assistant Professor Wai-Tat Fu, for all his support and guidance during my Master's study. From his intelligence and personality, I have learned a lot about how to be a good researcher. Also thanks to Adjunct Assistant Professor Anna T. Cianciolo, who has been providing important guidance and advice to me. I am also grateful to Peter Vlach, Karen Ayers of the Human Factors Division for their help, and my labmates (Jessie Chin, Michelle Moon, Wei Dong, etc.) for providing helpful suggestions to my study. Lastly, I hope to thank to my parents, Bin Kang and Xiaohong Wang, for their love and support.

# TABLE OF CONTENTS

<b>CHAPTER 1. INTRODUCTION .....</b>	<b>1</b>
<b>CHAPTER 2. THEORETICAL BACKGROUND .....</b>	<b>4</b>
2.1 THE INFLUENCE OF DOMAIN EXPERTISE ON INFORMATION BEHAVIOR .....	4
2.2 EXPLORATORY INFORMATION SEARCH AND SOCIAL TAGGING SYSTEMS .....	7
2.3 SEMANTIC IMITATION MODEL OF SOCIAL TAGGING SYSTEM .....	9
2.4 SUMMARY .....	13
<b>CHAPTER 3. RESEARCH QUESTIONS .....</b>	<b>14</b>
<b>CHAPTER 4. HYPOTHESES .....</b>	<b>16</b>
<b>CHAPTER 5. THE EXPERIMENT .....</b>	<b>19</b>
5.1 METHOD .....	19
5.2 RESULTS .....	23
5.3 SUMMARY .....	35
<b>CHAPTER 6. CONCLUSIONS.....</b>	<b>37</b>
<b>CHAPTER 7. DISCUSSION AND IMPLICATION.....</b>	<b>39</b>
<b>REFERENCES.....</b>	<b>42</b>
<b>TABLES AND FIGURES .....</b>	<b>46</b>

# CHAPTER 1. INTRODUCTION

As the World Wide Web is playing a more and more pivotal role in people's information and knowledge acquisition activities, the traditional way of searching for information on the Web has turned into a complex and multidimensional process. The information activities with more exploring, learning, and sharing are gradually replacing simple fact-retrieval activities. Therefore, theories about traditional information behavior are no longer sufficient for analyzing current user activities on line. Although more and more attention is being paid to the interactive information processes such as exploratory search (Marchionini, 2006) and information foraging (Pirolli, 2009), there is still a lack of research on how the individual's domain knowledge interacts with different information environments to influence users' information behavior.

The importance of individual's knowledge background has been explored in various research areas. In traditional information models about the cognitive process of information seeking (Russell, Stefik, Pirolli, & Card, 1993; Wilson, 1997), individual characteristics are proposed to have intervening effects on users' information behavior. Domain knowledge is assumed to have direct influence on the interpretation and learning processes during information search. There are empirical studies in consumer behavior showing that people with better knowledge will encode information and acquire information more efficiently (MacInnis and Jaworski, 1991). Search performance is also found to be better when users are searching in their own domain of expertise, as reflected by their query generation, website selection, search efficiency, and so on. Especially when novices and experts are inherently involved in the same social information environment, investigating the influence of domain expertise on information search not only will provide us with better understanding of people's search performance, but may also help to improve novices' learning and knowledge acquisition.

Today's prevalence of social information systems has brought individual information seekers into a more collaborated information network (Sen, et al., 2006). The popularity of collective intelligence platforms such as Wikipedia (<http://en.wikipedia.org>), Delicious (<http://delicious.com/>), CiteULike

(<http://www.citeulike.org/>) has stimulated intense discussions in the research community. The social environment provided by this kind of collaborative information systems is proposed to have the potential to support better information search (Millen, Yang, Whittaker, & Feinberg, 2007; Fu, 2008; Pirolli, 2009). For instance, tags created by users in the social bookmarking systems are supported to represent the semantic interpretation of the information content from other users, and therefore have the potential to play an active role in facilitating exchange of knowledge structures among users (Fu, 2008; Fu, Kannampallil, & Kang, 2009). Social tags, considered as trails for information search, would facilitate the exploratory process of learning and knowledge acquisition (White, Drucker, Marchionini, Hearst, & schraefel, 2007; Millen, et al., 2007). It is therefore reasonable to assume that, social information systems are more desirable than traditional search engines for exploratory search, as social tags can act as navigational cues that guide the iterative searching and learning process. Researchers have also argued that the collective intelligence provided by the social information systems are more useful for people to make sense of the information content (Marchionini, 2006; Fu, 2008). Recent studies (e.g., Kammerer, Nairn, Pirolli, Chi, 2009) have provided empirical support to the idea that social information systems can facilitate better search performance and learning effect for domain novices. However, there is still no direct evidence that shows specifically in what ways a social information system may assist exploratory search more effectively than traditional search engines, and how people's domain knowledge may play a role in the exploratory search process supported in social information systems.

To summarize, information behavior involved in social information systems has become more complex and dynamic than traditional fact-retrieval activities. Information seekers' domain expertise is believed to have the potential to guide the search and learning processes. However, there is still in lack of (1) empirical data to help us understand the different roles that domain expertise play in traditional and social search environments, and (2) how domain expertise and search interfaces work together to influence users' information behavior and knowledge acquisition.

In this thesis, I will first present a theoretical background review to summarize related research and our previous study on the topic learning and semantic imitation

model about social tagging systems. A laboratory-based experiment investigating the information seeking behavior in individual and social environments will then be introduced. The major questions addressed in this thesis are: (1) how experts and novices search for information differently in individual and social search environments; (2) how experts and novices interpret and learn information in the two environments differently; and (3) given that the individual characteristics would lead to differences in information processing, what are the implications of these differences to the design of future information systems.

## **CHAPTER 2. THEORETICAL BACKGROUND**

### **2.1 The influence of domain expertise on information behavior**

From the perspective of evolution theory, domain specificity originates from both social and inner space, as the external social world is proposed to impose its content to the internal. Psychologists (Cosmides & Tooby, 1994) have suggested that human perception and reasoning is guided by a collection of domain-specific systems of knowledge (Carey & Spelke, 1994). They also quoted the suggestion from optimal foraging theory that we should have domain-specific information-processing mechanisms governing foraging and sharing, and the mechanisms should be sensitive to different kinds of information in the foraging process (Cosmides & Tooby, 1992). On the other hand, theories about domain specificity have also been applied in the area of information science.

There has been a long history of research on how domain expertise may influence information seeking activities. Wilson (1981) described a general model of human information-seeking behavior, and discussed the factors that involve the context of information need, the information barriers that moderate behavior, and the actual information-seeking activities. He suggested that people's information need, considered as the root of the information seeking process, is driven by the information environment, users' social role, and users' psychological, affective and cognitive states. Domain knowledge, often measured by the level of education of the individual, is proposed to influence the interpretation and understanding of information. Wilson provided examples to show that people with higher domain knowledge have less motivation to seek for more information on the same topic, but people with less domain knowledge will tend to search more. From the perspective of traditional information science, Wilson was one of the first researchers who provided broad-brush theories to characterize the influence of domain knowledge on information need and information seeking behavior.

Studies on sensemaking also revealed some relationship between individual knowledge representation and the information foraging process. Sensemaking is



considered as the process of “how people make sense of the external world”. One important branch of sensemaking studies is in information science, focusing on how people process and organize the information. Russell, et al. (1993) used the example of laser printers to explain that the sensemaking process includes cyclic processes of searching for representations and encoding the information, referred as the “learning loop complex”. This learning loop includes four major steps: search for representation, instantiate representation, shift representation and consume encodons. They also claimed that, “Sensemaking is the process of finding a representation that organizes information to reduce the cost of an operation in an information task”. The original knowledge representations (or schemas) in people’s mind could be *expanded, merged, split* or *added* if the external knowledge representations do not fit the established categories, so individuals may merge or edit the “categories” of information to fit their established structure. The external knowledge representations, according to Pirolli (2009), could provide knowledge to help people more adaptively engage their task environment. It is also a potential source of valuable knowledge to improve people’s ability to accomplish their goals. Therefore, the match between individual’s knowledge representation and the structure of external environment will be influencing how people process the information and conduct related tasks.

After reviewing several commonly adopted information processing models, one can see that domain expertise is playing an important role in guiding users’ information behavior. Further, researchers also provided some empirical evidence to support this assumption. Bhavnani (2001) examined how experts in health care and online shopping search for information within and outside their domains of expertise. They distinguished the domain-specific search knowledge into two parts. The declarative components consist of knowledge about classes of websites within a domain, knowledge of specific websites (such as their URLs), and content knowledge consisting of the nature and type of information within a website. The procedural components consist of sequencing knowledge that allowed them to formulate an overall search plan based on their conception of the different classes of websites, as well as termination knowledge that allow them to decide when to end a search. He found that when performing tasks within

the user's domains of expertise, he or she would use declarative and procedural components of domain-specific search knowledge to perform effective searches. In contrast, when they performed tasks outside their domains of expertise, they used a range of general-purpose search methods that lead to comparatively less effective search results.

Duggan & Payne's study (2008) has shown that better domain knowledge led to less time spent on each webpage, faster decision to abandon inquiry, and shorter queries being entered into search engine. They also found that domain knowledge could affect the quality of queries users entered into the search engine. In addition, users' background knowledge would increase their ability to select links that would more likely lead to the target information. White, et al. (2009) summarized the influence of domain expertise on user's web search behavior by conducting a log-based analysis on a large-scale data. They investigated users' queries, search sessions, website selection, and rates of search successes for domain experts and non-experts. They found that domain experts were more successful in search and used more domain specific vocabularies in their queries. Allen (1991) examined the impact of topic knowledge on information catalog searching. They found that higher-knowledge participants used more search expressions in catalog search than lower-knowledge participants. They also suggested a possible trend that participants with higher level of topic knowledge and who expressed difficulty in search were more likely to introduce new vocabulary into the search. Hsieh-Yee (1993) found that when users had certain levels of search experience, domain knowledge would play an important role affecting the reliance on their own language and the use of the external information content as search queries. The result indicated that when users were searching out of their domain, they made more effort in preparing for the search, monitored the search more closely, and tried out more term combinations; but when experts searched in their domain, they used more of their own terms to search. Similar to the results from these studies, Zhang, Anghelescu, & Yuan (2005) found that experts generated more queries in search and more words in each queries, but the search efficiency did not show any difference between experts and non-experts.

Considering information search as a cognitive process, Rouet (2003) proposed that information search involves both reasoning and memory search processes. He conducted an experiment with university students in specific majors to see whether domain knowledge would influence search strategies. Interestingly, their experiment showed that students' strategies depend more on search questions but less on students' prior knowledge about the domain. When asked specific questions, students in two majors performed fast and precise searches, with very few lookbacks to the question. In contrast, when asked general questions, all students conducted longer searches and looked back more often to the question. The "general questions" in this experiment are very similar to the concept of exploratory search (Marchionini, 2006), which I will discuss in the next subsection.

In general, most of the existing studies focused on how domain expertise affects the search related performance data, such as query generation, website selections, search strategy, search efficiency and so on. However, few studies have mentioned the differences between domain experts and novices while they are using *collaborative information systems*, which, as I will elaborate in the next chapter, will be one of the major research questions in this thesis.

## **2.2 Exploratory information search and social tagging systems**

Recently, researchers have become more and more interested in information seeking processes that are exploratory in nature. There are often situations in which the information seeker has not yet developed well-defined information goals to guide their search. Instead, the information seeker may have to start with an abstract representation of information needs derived from a broader task context. In these situations, the information seeker has to engage in some forms of *exploratory information search*, through which information goals can be iteratively refined and enriched (e.g., Fu, 2008). Recently, researchers have reasoned that the traditional search engines might be insufficient for this kind of exploratory search (Marchionini, 2006). Instead, many have proposed that the evolving Web 2.0 technologies have greater potential for helping people to conduct exploratory information search. Social bookmarking systems (or social

tagging systems) have been discussed a lot for its usage on facilitating this kind of exploratory information seeking.

Social tagging systems allow users to annotate, categorize and share web content (links, papers, books, blogs, etc.) using short textual labels called tags. The inherent simplicity in organizing and annotating content in these systems through “open-ended” tags satisfies a personal and social function (Ames & Naaman, 2007; Thom-Santelli, Muller, & Millen, 2008). At a personal level, customized tags can be added to a resource based on personal understanding and individual purposes that will help in the organization of resources or for future search and retrieval. At the social level, tags can facilitate sharing and collaborative indexing of information, such that social tags act as “way-finders” for other users with similar interests to search for relevant information (Fu, 2008; Kammerer, et al., 2009; Kang, Fu, & Kannampallil, 2010; Millen, et al., 2007; Piroli, 2009). More recently, a number of studies have explored the potential of social tagging systems on helping improve the search performance. Morrison (2008) argued that social tags, interpreted as folksonomies, would have as much precision as search engine. He also suggested that folksonomies may facilitate the finding of new information compared to search engines. Heymann, Koutrika, & Garcia-Molina (2008) analyzed large datasets from Delicious and suggested several interesting relations between tags and URLs. Similar to results obtained by Morrison, they found that users in Delicious were more interested in newly added URLs. This recency effect, similar to Cattuto’s work (Cattuto, Loreto, & Pietronero, 2007), may be attributed to the possibility that those URLs created recently are tagged more frequently. They also found a relatively high overlap between popular query terms and popular tags. They therefore argued that most tags in social bookmarking system are relevant and effective for information search. They also proposed that tags associated with bookmarks in a bookmarking system are more useful than typical link texts (e.g., page titles) returned from a search engine. Krause, Hotho, & Stumme (2008) compared user activity and behavior from Delicious, MSN, AOL and Google by analyzing tags and queries. By comparing the total number of query terms in MSN and tags in Delicious, they showed that MSN has a significantly larger number of terms but the average frequency of each item was quite similar in both

systems, indicating that Delicious users focused on fewer topics, but each topic was reached by users equally often. All of these studies include some empirical data supporting that social tagging systems might be able to provide a better information search environment with the presence of human-generated indices that facilitate exploratory search.

The social influence among users in a social tagging system is also considered an unique feature that is helpful for understanding the exploratory search process. Previous studies suggested that users could benefit from a social search environment by reading information cues (e.g., social tags) left by other users, that act as “signposts”, guiding them to the right information (e.g., Heymann, et al., 2008). By using these tags as trails for information search would lead users through an exploratory process of learning and knowledge acquisition (White, et al., 2007). Golder & Huberman (2006) found that users’ tag choices were influenced by others’ tags even if they had different information needs when tagging. Sen, et al. (2006) showed that available tags in a tag community could directly impact a user’s tendency in choosing tag vocabulary. Most recently, our work (Kang, Kannampallil, He, & Fu, 2009) suggested that the social environment in tagging systems, as suggested in these prior research reports, would have the potential to support knowledge exchange during the information search process.

To sum up, information *searching*, *exploring* and *learning* are three major steps in exploratory information search (Marchionini, 2006). The *sharing* of information is also an important part of users’ behavior in social information systems. Although there are separate studies focusing on different parts of these steps, none of them has discussed how domain expertise would affect information processing in a social context, and whether the effects of domain expertise would differ between individual and social environment.

### **2.3 Semantic imitation model of social tagging system**

Besides the comparison between social tagging systems and traditional search engine mentioned in the previous part, another intriguing feature of social tagging systems is that they can be considered platforms for dynamic interactions of diverse

semantic structures among users (Cattuto, et al., 2007). If features of social tagging systems can influence higher-level knowledge structures of users, social tags not only may provide annotation to web contents, but they may also have the potential to play an active role in facilitating exchange of knowledge structures among users (Fu, 2008; Fu et al., 2009). By looking at the tags created by other users, people can develop their own interpretation of the information based on the cues from social tags and URLs.

Referring to the research about reading comprehension and information extraction, as a person reads text, words invoke corresponding semantic representations to allow the person to extract meaningful information contained in the text (Kintsch, 1998). This kind of spontaneous semantic interpretation of words is perhaps best illustrated by the experiments on “false memories” (Roediger & McDermott, 1995). A typical false memory experiment would show that when people were asked to remember a list of semantically associated words that converged on a non-studied word, people tended to falsely remember the non-studied word. For example, after studying the list consisting of thread, pin, eye, sewing, sharp, point, pricked, thimble, haystack, pain, hurt, and injection, people often erroneously recalled the converging non-studied word needle in the list. This kind of “memory illusion” is often interpreted as evidence supporting the notion that as people process a list of words (or tags, when they are browsing a social tagging system), they spontaneously activate the corresponding semantic representations for those words. When people try to recall the list of words, the converged semantic representation will again be activated to exert a top-down influence on memory recall. As the false-memory experiments showed, because the non-studied word was representative of the converged semantic representation, it was often erroneously “recalled”.

Results from these experiments therefore demonstrated that people tend to naturally encode *semantic representations* of words during comprehension. Derived from these theories, the semantic imitation model (Fu, et al., 2009) decomposed the social tagging process into two parts: a topic inference process and a topic extraction process (as shown in Figure 1). As the information seekers navigate through a social tagging system, tags created by other users will help them interpret whether a particular piece of information would be relevant to search goal. The set of tags assigned to the bookmark

will act as retrieval cues for relevant topics (or concepts) represented by these tags. This is called the tag-based topic inference process. Thus, the process assumes that the topics inferred from the tags will allow the user to predict the information content of the associated resource as well as to provide some form of semantic priming of related concepts when the user processes (comprehends) the information in the resource (Kintsch, 1998). The topic extraction assumed that the user extracts the concepts (topics) that describe the contents of the document, influenced (i.e., biased) by the initial tag-based topic inference (Griffiths, Steyvers, & Tenenbaum, 2007). The model assumes that when a user processes a resource, he or she will engage in a process of topic extraction to comprehend the associated information content (Fu, 2008; Pirolli, 2004). Associated information content can include abstracts of papers (in CiteULike) or overviews of web URLs (Delicious), or the complete content of a web page. As the iterative topic inference and extraction happen in the searching and tagging process, it is reasonable to believe that it can also facilitate users' learning about the search topic.

More specifically, the model was used to examine the social influences in social tagging systems. Two sets of simulated users were created with differences in their background knowledge structures – domain experts who had perfectly matched word-concept distribution with the documents created; and domain novices who had a less well-structured knowledge representation. As shown in the top part of Figure 2, experts reached stability much faster than novices. The faster convergence in the case of experts could be explained by the fact that tags assigned to each document were more predictive of the topics contained in the document, and that the experts were much better at extracting the correct concepts based on “high quality” tags created by other experts. The bottom panels in the figure show the scatter-plots of the relative tag frequencies of one special document that we created to illustrate this difference. This special document contained a single topic, with the mean of the prior distribution of words over this single topic at word 300. As expected, for both experts and novices, tag proportions were highest around the most representative words. However, experts clearly had a much more focused vocabulary than novices, as shown by the narrower spread of tag choices. In addition, novices seemed to have “misinterpreted” the topic and chose tags around word

800 (the initial choice of this tag was due to random noise) to represent the wrong topic. This simulation result provides theoretical background of the differences between domain experts and novices in interpreting documents and assigning tags, showing that domain experts have more predictive and more converged tag choices than novices. Our previous study has also provided some empirical supporting the semantic exchange between users in social tagging system (Kang, et al., 2009).

In that study, participants were invited to conduct information search on CiteULike for assigned information tasks, and classified them into social and individual search groups. Participants were allowed to see other people's tags in the social groups, but not able to see others' tags in the individual group. Participants in each group were organized into 4 sessions, and tags created in each session were accumulated across time. That study controlled the information environment and search tasks, in order to analyze the interaction effect between search tasks and social influences on the tagging process. The number of tags created in each document were calculated, and the semantic relationship between tags was analyzed using Latent Semantic Analysis (Laudauer & Dumais, 1997). As shown in Figure 3, the linear downward trend for number of unique tags across sessions for the social group was significant, suggesting that as more tags were added to the library, the number of unique tags decreased, but the individual group did not show any significant trend. The LSA results in Figure 4 further validated the social influence on the semantic level. It is obvious that in the social group, the LSA scores stayed approximately at the same level for tags created under the same information goal across sessions, but the LSA scores increased significantly across sessions for tags created under different information goals. The LSA scores between tags in the individual group were very low in both conditions (same and different information goals). In other words, the influence of social tags eventually outweighed the influence brought by difference information goals and caused the semantic convergence of tag choices.

Results from the above empirical experiments validated the assumption about the social influences of tags on semantic level in the semantic imitation model, but there is still no empirical data supporting the expertise difference proposed by the model. For instance, the unique tags in Figure 3 decreased, but what indeed caused this pattern, and



whether this pattern would differ if the information seekers have expertise knowledge? Also, as tags have the potential to support the knowledge exchange among users, a natural question is how domain experts and novices would have different performance when using social tagging systems, whether the social tags would have different social influence on the two kinds of users, and how can we provide future implications to facilitate the knowledge exchange and learning. Although there are lots of studies validating domain expertise' influence on information processing, none of them has investigated how users interpret the information and users' learning effect evoked by social tags.

## **2.4 Summary**

In this chapter, I introduced several information models explaining the influence of domain expertise on information processing, reviewed relevant literatures about exploratory search and social tagging systems, and briefly presented previous work on a semantic imitation model and empirical data about social tagging systems. Previous research has provided sufficient support in distinguishing the influence of domain expertise on information search. Researchers have also pointed out the potential usefulness of social tagging systems for supporting exploratory search by both experts and novices. Although there are many comparisons between social tagging systems and traditional search engines, as well as studies about the effect of domain expertise in information search, there are fewer studies about how domain knowledge impacts users' search behavior in different types of search interfaces. What is still missing is how the expertise profiles of users affect their exploratory information search and knowledge learning in a social search system.

## CHAPTER 3. RESEARCH QUESTIONS

As Wilson (1991) suggested, individual characteristics such as domain expertise would influence information processing in several aspects. Domain expertise may also play a role in directing the shift of knowledge representation while users are making sense of new information from the environment (Russell, et al., 1993). Researchers showed that greater domain knowledge would lead to higher quality of information acquired, better search efficiency and better query generations (Duggen & Payne, 2008; White, et al., 2009; Fu, et al., 2009). Domain experts were also found to make more use of their own knowledge to search, when compared with novices in the same environment (Allen, 1991; Hsieh-Yee, 1993). As the traditional search engines are suggested to be insufficient to satisfy users' growing information demands, social tagging systems are believed to have the potential to facilitate exploratory information search (Marchionini, 2009; White, et al., 2007; Kammerer, et al., 2009). Researchers also suggested the possible usage of social tags on supporting knowledge exchange and semantic interpretation (Cattuto, et al., 2007; Fu, et al., 2009), and discussed the possible differences between domain experts and novices in conducting exploratory search on social tagging systems (Fu, et al., 2009; Kammerer, et al., 2009).

However, none of the previous studies has provided empirical answers to the question of whether domain expertise would induce different understanding to topic-related information, and whether it would influence information behavior differently in a social environment. Attempting to address this problem, this thesis aims at answering the following four research questions:

RQ 1. Would domain expertise influence how users conduct *exploratory* information search? What is the role of domain knowledge in the steps of information *seeking, learning* and *sharing*?

RQ 2. Will domain expertise influence the information collected by users (e.g., different topics, or different types of information)? Will domain expertise also influence how users interpret the information about the same topic? Is there any

learning effect in domain experts vs. novices afforded by different search environments?

RQ 3. If domain experts and novices have difference performance while searching for the same topic, would search interfaces affect their search behavior? How would search environment and domain expertise work together to drive the exploratory information search?

RQ 4. Given the assumption that domain expertise and search interfaces would cause differences in search behavior, what implication can be drawn on the design of future information systems? How can we improve the collaboration between different user groups from the system?

A laboratory-based experiment was designed to test the above research questions. Subjects with different level of knowledge in the same domain were recruited to conduct information search in controlled environments. The experimental results including websites collected, tags attached, knowledge tests, and interviews were analyzed to help answering these research questions.

## CHAPTER 4. HYPOTHESES

Based on the above general research questions, I formulated the following testable hypotheses to provide answers to the questions. The first hypothesis focuses on the behavioral difference between experts and novices while they are searching in different environments. The goal is to answer the first research question: will domain expertise influence the exploratory information search? The null hypothesis for this question is that domain experts and novices will have similar behavioral pattern in two search environments. As I will describe in the next chapter, an experiment was conducted, and search queries, URL visits, bookmarks saved and tag created were recorded to test this hypothesis. By comparing their behavioral data, I expect to find different patterns in experts and novices to reject the null hypothesis.

Hypothesis 1. Domain expertise will induce different search behavior in exploratory information search.

- (a) Domain experts will find more useful information from the social web, and also contribute more shared contents (tags) to the social web than domain novices.
- (b) Domain experts will adopt more knowledge-driven search strategies, and use more of their own queries to search for information. Domain novices will adopt more interface-driven search strategies, and utilize the social cues (e.g., tags) provided by the systems more.

The next hypothesis focuses on the learning and interpretation process in the exploratory search. Given that domain experts may have richer concepts and more complex knowledge structures in their head, the way that they interpret the information content might be different from domain novices. As a result, the learning effect of information search on experts and novices might also be different. The null hypothesis for this question is that (a) tags (assumed as the interpretation to the information content) created by domain experts and novices will be similar; (b) both domain experts and domain novices will gain similar level of knowledge from the search process, and the

information gain brought by two environments will not have difference. Participants' tags will be analyzed to examine their interpretation to the information, and some knowledge tests will be used to reflect their knowledge change.

Hypothesis 2. Domain experts and novices will interpret information differently, and gain different level of knowledge from the web.

- (a) Domain experts will be able to find more specific information as well as general information, while domain novices will mostly find general information.
- (b) Domain experts will have similar interpretation (tags) with each other, but domain novices will have more diverse interpretation.
- (c) After the information search, domain novices will have more knowledge change than domain experts.

Besides the influence of domain expertise, this thesis also examines the influence of search environments on search behavior, and discusses the interaction effect between domain expertise and search environments. The null hypothesis corresponding to the third research question is that the search behavioral data will not show differences between individual and social search environments. The search behavior of participants using the two interfaces will be analyzed to test the hypothesis. The expected result is stated below.

Hypothesis 3. Domain expertise will play different roles in individual and social search environments.

- (a) Novice will find more information in social environment compared to individual environment, but domain experts will find similar amount of information in both environments.
- (b) Domain experts will have more shared information content in social information system, but they will find more unique concepts in individual system.
- (c) Social information systems will facilitate better learning effect.

The fourth research question is an open-ended question with no specific hypothesis set before the study. Implications will be discussed in Chapter 6.

## CHAPTER 5. THE EXPERIMENT

A laboratory-based experiment was conducted to test how experts and novices perform exploratory information search in two different search contexts: individual and social search environments. I measured the behavioral data in information search, and analyzed the information they collected and how they interpret the information. The purpose is to investigate how domain expertise influence users' search behavior in different search environments. Under the assumption that domain expertise would influence how well participants could generate keywords to search, interpret the search results, and select social tags, I also expected that experts and novices could adopt different search strategies and gain information differently when they performed exploratory search using the two interfaces.

### 5.1 Method

A  $2 \times 2$  between-subject design was used to investigate the differences in users' search behavior when they were using a traditional search engine (Google) and a social tagging system (Delicious), and how users with different levels of domain expertise would interpret information differently using these two different search interfaces (as shown in Table 1. While Google provides a traditional search environment for keyword-based queries, Delicious provides tagged social bookmarks created by other users that allow participants to use either tag-based or keyword-based queries to search (as shown in Figure 5).

#### 5.1.1 Participant

A total of 48 participants were recruited for the study (22 female, 26 male, mean age = 24.4). All participants were skilled computer users with more than 10 years of computer usage experience (mean = 13.8). All of the participants reported Google as their most familiar search engine and that they performed Internet searches with an average frequency of 3.95 on a 5-point scale (4 means "use search engine several times a day"). 24 of the participants claimed to have expert knowledge in finance or related area (such as holding an advanced degree or had current or prior employment experiences in the

finance industry). The other 24 did not have any training or knowledge in finance or related fields. In addition to the self-claims on their knowledge backgrounds, three additional methods were used to further verify participants' expertise level (discussed later). Expert and novice participants were randomly assigned to one of the two interfaces in the  $2 \times 2$  experimental design. Participants were paid \$25 for their participation in the experiment.

### **5.1.2 The Exploratory Search Task**

“Financial crisis” was used as the topic for the exploratory search task. This topic was chosen for its current relevancy and differences in the depth of knowledge about the topic between subject matter experts and the general public (topic novices). Participants were asked to imagine that they were to collect information from the Web to give a talk on the current financial crisis. They were encouraged to explore information using their assigned search interfaces (Google or Delicious) to enrich or supplement their own knowledge. During their search activities, participants were asked to save and tag useful websites as bookmarks. In Delicious, they could save websites as bookmarks to their assigned web account, while in Google they were instructed to save bookmarks in a given folder in browser and create tags for the resource. They were instructed to search, read, and select information, but not spend too excessive time on a single web page. The following data was collected from each participant.

#### *Self-report*

Participants were asked to complete a short survey with 5 questions about their knowledge of finance and economics as well as their familiarity of the current financial crisis on a 5-point scale. Sample questions include: “I know the causes and backgrounds of the current financial crisis”, “I can give my own opinion about what should be done to deal with this crisis”, etc. We found a high reliability for the self-report questionnaire (Cronbach's alpha = 0.921).

#### *Knowledge Questionnaire*



A knowledge questionnaire was used to test the participants' specific knowledge about the financial crisis. There were 20 questions in total, of which 10 were general questions such as, "Which event precipitated the current financial crisis?" The rest 10 questions are specific questions that required professional training in finance or economics (e.g., "Which of the following is the investors' strategy against the unsystematic risk?"). Questions in the knowledge questionnaire were collected from online quizzes about the financial crisis and from textbooks. The questionnaire was reviewed by two graduate students majoring in finance and one professional with more than 15 years experience in a financial holding company.

### *Topic description*

The participants were also asked to perform a topic description task before and after they did the information search. In this task, the participants were given the topic "financial crisis" and were asked to write down words or phrases to describe the topic. This task tested the fluency of the concepts that the user generated to associate with the topic. The purpose was to measure their understanding about the topic based on their retrieval of terms and concepts from the memory (Griffiths, et al. 2007). One would expect to see a conceptual knowledge change by analyzing the topic description results before and after information search. In addition, as proposed in the semantic imitation model (Fu, et al. 2009), there is a topic extraction process in tagging. I therefore also expect to validate the relatedness between tags and topics gained from the information search.

### *Categorizing*

Bookmarks and tags created by the participants were presented to them after the information search task. The participants were required to categorize their bookmarks and provide a label to each category. The categorization of bookmarks is considered as a direct measure of whether the selection of bookmarks is influenced by their domain knowledge. This categorization task was designed to help to examine participants' knowledge gain on a higher level and see how experts and novices interpret the search results differently based on their own knowledge structures.

### **5.1.3 Procedure**

Participants were first given general information about the experiment and the goal of this study, and were then asked to read and sign the consent form for participating in the experiment. Participants then filled out a general survey about their demographics and a self-report survey on their knowledge background. After that, they were asked to do the pre-test topic description task. For the topic description task, they were asked to write down terms/phrases about the topic on a sheet of paper and stop at anytime when they were done. On average, the topic description task took about 5 minutes. Then participants were randomly assigned to the Google or Delicious condition. The researcher briefly explained the task and demonstrated how to use the search engine or the social tagging system and how to create tags and save bookmarks. Participants were provided enough time to familiarize themselves with their tasks and the interfaces before they started the experiment, during which the experimenter would answer any questions that they had. Participants performed their tasks individually and were given a maximum of 1.5 hours for their task.

The Camtasia recorder was used to record all on-screen actions of the participants including information searching, bookmark selection, tag creation and URL clicks. After finishing the search task, participants performed a post-test topic description task. Then they completed the knowledge questionnaire. The knowledge questionnaire was given after the search task to avoid potential priming effect on their search behavior by the knowledge questions. After the participants filled out the knowledge questionnaire, the researcher provided a printed copy of the bookmarks and tags that they had generated during their search task. The participant was then asked to categorize the bookmarks into groups and give a label to each group. A short open-ended interview was conducted in the end regarding the participants' opinions about the search interface, tagging process and categorization. A brief flowchart of the procedure is shown in Figure 6. The whole experiment took about 2.5 hours.

## 5.2 Results

### 5.2.1 Identifying Domain Knowledge

From the self-reported expertise ratings, a significant difference was found between the groups (mean = 3.8 and 2.87 on a 5-point scale for experts and novices respectively,  $p < 0.001$ ). Consistent with the self-reported ratings, there was also a significant difference on the general knowledge test score between experts and novices ( $p < 0.05$ ), as well as on the 10 professional questions in the questionnaire ( $p < 0.01$ ). Experts also generated more terms to describe the topic of "financial crisis" than novices, both before and after the task, but marginally significant. All these tests validated the assumption about experts' higher domain knowledge than novices. Detailed statistics could be found in Table 2.

### 5.2.2 Search Behavior

To analyze the differences in their search behavior, I compared the number of bookmarks and tags created, the number of URL visits, and the number of URL visits per bookmark saved for each participant across the groups. While the number of bookmarks and tags created could reflect the effectiveness of their search behavior, the number of URL visits per bookmark saved could indicate how efficiently participants could find relevant information using the interfaces.

As Figure 7.a shows, the analysis of variance (ANOVA) showed that experts saved more bookmarks than novices ( $F(1, 44) = 2.52, p = 0.1$ ), but the interaction effect between expertise and interface was not significant ( $p = 0.2$ ). Post-hoc analysis showed that experts collected more bookmarks in Delicious ( $p < 0.001$ ), and the difference between experts and novices in Google was not significant. As shown in Figure 7.b, the main effect of interface was significant for the total number of tags created ( $F(1, 44) = 4.105, p < 0.05$ ), indicating that participants using Delicious generally had higher number of tags than participants using Google (mean = 79.2 and 59.1). The interaction effect of expertise and interface was also significant for the number of tags created ( $F(1, 44) = 6.146, p < 0.05$ ). The post-hoc analysis showed that experts using Delicious generated

more tags than novices in Delicious ( $p < 0.05$ ), but the difference between experts and novices was not significant in the Google group.

As expected, domain expertise facilitated information search in both Google and Delicious, as reflected by the higher number of bookmarks saved by experts. Experts also created significantly more number of tags when using Delicious than novices. This can probably be attributed to the fact that experts were better at interpreting tags created by others, as well as generating terms to describe the bookmarks (I will present further analysis on this). However, domain expertise did not induce the same difference in the Google group. At least in terms of the total number of bookmarks saved and the number of tags, novices and experts were about the same in their performance when using Google.

Figure 7.c shows that experts visited more URLs than novices in both interfaces, although the difference was only marginally significant ( $p = 0.14$ ). The main effect of interface and the interaction between interface and expertise was not significant for number of URL visits. However, when analyzing the number of URL visits per bookmark saved (see Figure 7.d), I found a significant interaction between interface and expertise ( $F(1,40) = 10.148, p < 0.01$ ). This measure of the number of URL visits per bookmark saved could partially reflect the efficiency of search, as a smaller number would indicate that the number of relevant bookmarks saved was higher per unit browsing action. Post-hoc analysis confirmed that the search efficiency for experts in the Delicious group was significantly better than novices ( $p < 0.01$ ). Interestingly, novices visited more URLs to find a relevant bookmark when they used Delicious compared to Google ( $p < 0.05$ ), which could indicate that novices had lower efficiency when searching in Delicious than Google.

The query generation of the four groups of participants was then analyzed. Figure 8.a shows the number of keyword-based queries (entering keywords in keyword search box) performed by experts and novices in each interface. ANOVA showed that the main effect of interface was significant ( $F(1, 41) = 7.341, p < 0.01$ ), as well as for the interaction between expertise and interface ( $F(1,41) = 3.109, p < 0.1$ ). The main effect of

expertise was not significant. Participants in Delicious generally used less keyword-based search than Google, and the interaction was mostly caused by the difference between experts and novices in the Delicious group. Indeed, post-hoc analysis showed that experts performed significantly more keyword-based queries than novices in Delicious ( $p < 0.01$ ), but the difference was not significant for the Google group. To further understand the difference in search strategy brought by expertise difference in the Delicious group, a separate ANOVA was performed on the use of tag-based queries (selecting tags from the popular tags or other users' tags attached to each website title) and keyword-based queries (entering keywords in keyword search box) for experts and novices in Delicious (as in Google there was no tag-based query). As Figure 8.b shows, the interaction between query type (tag-based or keyword-based) and expertise was significant ( $p < 0.05$ ). Post-hoc analysis showed that experts used more keyword-based search than novices ( $p < 0.05$ ), while novices used more tag-based queries than experts.

To summarize, this part of results suggest that experts and novices search differently in two search environments, in terms of the number of bookmarks saved, tag creation, search efficiency in finding relevant information, and the search strategies. In this experiment, experts collected more information (as reflected by the number of bookmarks), created more tags, and have higher search efficiency in Delicious compared to novices in the same condition (Hypothesis 1a). Domain expertise was found to be a major factor influencing the information collection. And the higher search efficiency in Delicious might be attributed to the fact that social tags may facilitate the evaluation of the relevance of links before users click on them. The interaction effect between expertise and query types in Delicious provided direct evidence that experts were more able to come up with their own queries to search (Hypothesis 1b). Experts conducted more keyword-search but less tag-search in Delicious. This result implied that experts used their own terms more (Hsieh-Yee, 1993) by using more keyword-based search, but novices relied more on directly using others' tags to search.

### **5.2.3 Bookmark Selection and Categorization**

In addition to search behavior, the information content (bookmarks) saved by different groups were also analyzed. 48 participants selected 1170 bookmarks. Among

those 1170 bookmarks, 359 bookmarks were saved by more than 2 participants independently, 811 bookmarks were saved by only one participant. In total 937 distinct bookmarks were saved by all users.

Figure 9 shows the frequency-rank plot of the bookmarks saved by all users. Consistent with previous studies, the plot of the number of participants sharing each bookmark shows a typical power-law distribution (Cattuto, et al., 2007), in which there are rapid drops in the frequencies of bookmarks as rank increased, and it also has a long tail indicating many unique bookmarks saved by individual participants. The most popular bookmark saved by 11 participants was the wikipedia page on subprime mortgage crisis ([http://en.wikipedia.org/wiki/Subprime\\_mortgage\\_crisis](http://en.wikipedia.org/wiki/Subprime_mortgage_crisis)).

#### *Popular vs. unique bookmarks*

To study the extent to which different interfaces and expertise may lead to the saving of more unique or more shared bookmarks, the bookmarks saved by all participants were divided into two groups: the bookmarks shared by 3 or more participants in our experiment were called popular bookmarks; the bookmarks shared by 2 or less participants in our experiment were called unique bookmarks. Each of the popular and unique bookmarks were shared by an average of 307 and 21 users in Delicious respectively, which at least partially validated the general "popularity" of these bookmarks as reflected by the massive number of users in Delicious. In order to find out how participants with different level of domain expertise and different interfaces would save bookmarks in the popular or unique groups, a 2 (shared frequency)  $\times$  2 (interface)  $\times$  2 (expertise) ANOVA was performed using the number of bookmarks saved by each user as dependent variables.

Results showed that the main effects of expertise and shared frequency (popular/unique) were significant ( $p < 0.05$ ), but the main effect of interface was not significant ( $p = 0.91$ ). The interaction effect of interface  $\times$  shared frequency and expertise  $\times$  shared frequency was significant ( $p < 0.10$ ). The interaction effect of expertise  $\times$  interface was marginally significant ( $p = 0.10$ ). The three-way interaction of interface  $\times$  expertise  $\times$  shared frequency was not significant.

As shown in Figure 10, both experts and novices selected more popular bookmarks when using Delicious than when using Google. For unique bookmarks, experts selected almost the same number of bookmarks in Delicious and Google. However, novices selected more unique bookmarks in Google than in Delicious. As presented in 5.2.2, the number of bookmarks collected by experts in Delicious was significantly more than novices (see Figure 6.a). Now we can have a clearer picture about the difference between experts and novices in Delicious. Apparently, the higher number of bookmarks created by experts than novices in Delicious was caused by the higher number of unique bookmarks selected by experts, as evidenced by the fact that the difference between experts and novices was significant in unique bookmarks ( $p < 0.05$ ), but not in popular bookmarks ( $p = 0.26$ ). This result gives strong evidence to the fact that experts relied more on their own background knowledge and were influenced less by the social environment in Delicious; but novices selected less unique bookmarks in Delicious because of the stronger social influence in that condition.

By examining the content of the bookmarks, it is found that most of the unique bookmarks were either specific web sites describing a particular event, or professional websites developed for finance professionals (White, et al., 2009). Therefore, it seemed that in human-generated indexing systems such as Delicious, novices were less likely to benefit from tags that experts gave on those unique websites, as novices might not have the background knowledge to judge whether or not those tags and websites were relevant information.

To further understand how domain expertise contributes to the differences in the selection of bookmarks, I analyzed how users categorized their own bookmarks based on the content (e.g., bookmark A might be categorized into the cause of the financial crisis, and bookmark B might be in the category about the history of the financial crisis). 242 categories were generated by 47 participants (one participant's data was lost due to technical problems). On average, each participant generated 5.15 categories, and the average number of bookmarks in each category was 5.18 ( $SD = 2.45$ ). These categories would help understand how experts and novices differed in their interpretation of the contents of the web sites they found, and why they believed the information was relevant

to the topic. Although users were free to select any words to categorize the bookmarks, that there were overlapping and semantically similar categories. Two raters merged the categories into 77 distinct categories by combining identical or semantically similar categories. The agreement between the two raters was 91%.

#### *Popular vs. unique bookmark categories*

Similar to the analysis of bookmarks, I classified the categories into two groups. In the popular group, more than three participants used the same category, and the rest were put in the unique group. The most shared category is “Cause of the financial crisis”, which was used by 11 users. Under this category, there were 154 distinct bookmarks. The popular categories were all general or common categories like “causes”, “history” or “explanation of the financial crisis”. In contrast to the highly shared categories, there were more categories that were unique to individual participants or shared by only two participants. Most of these categories were related to a specific company, person, event, or professional terms (e.g. “AIG”, “CDO”, etc). Table 3 shows a list of sample categories. In order to further understand how users interpreted the information they collected, I hope to find out whether participants with different level of domain expertise using different interfaces might be more likely save bookmarks in the popular or unique category groups. Similar to the analysis of bookmark selection, a 2 (shared frequency) × 2 (interface) × 2 (expertise) ANOVA was performed using the number of bookmarks in each category as dependent variables.

Results from ANOVA showed that the main effects of shared frequency and interface were significant ( $F(1, 367)=2491.5, p<0.001$  and  $F(1, 367)=100.4, p<0.05$ ), but the main effect of expertise was not significant ( $p=0.26$ ). The three way interaction of expertise × interface × shared frequency was significant ( $F(1,367)=73.7, p<0.05$ ). This three-way interaction was apparently due to the significant two-way interactions between interface and shared frequency ( $F(1, 367)=242.1, p<0.001$ ), as no other 2-way interaction was significant.

As the main effect of expertise was not significant, separate ANOVAs were performed on each level of expertise. For the expert group, the main effects of interface



and shared frequency were significant ( $F(1, 217)=1162.8, p < 0.001$  and  $F(1, 217)=78.0, p < 0.05$ ). The interaction of interface  $\times$  shared frequency was also significant ( $F(1, 217)=313.1, p < 0.001$ ). For novices, only the main effect of shared frequency was significant ( $p < 0.001$ ), but all other effects were not significant. As shown in Figure 11.a and b, experts saved more bookmarks in popular categories when using Delicious than when using Google ( $p < 0.001$ ) but the reverse was true when they are selecting bookmarks that belong to the unique categories ( $p < 0.001$ ). Novices also selected more popular bookmarks in Delicious than in Google ( $p < 0.001$ ), but the difference between Delicious and Google was not significant when novices were selecting bookmarks in unique categories. In addition, post-hoc analysis was conducted to compare the differences between experts and novices in four cases, and it is found that only the difference between experts' and novices' bookmark number in unique categories differed significantly in Google condition ( $p < 0.001$ ) as shown in Figure 11.b. All other comparisons were not significant.

This pattern of results suggests that for both experts and novices, when they used Delicious to search for information, they saved more general information than using Google. Interestingly, when experts used Google, they actually categorized more information into unique concepts than when they used Delicious, but novices found less number of unique information in Delicious and Google. Compared with the bookmark results, we can see that although experts selected almost the same number of information that is less popular in two interfaces (Figure 10.b), they still found more bookmarks with unique concepts when using Google (Figure 11.b). The results were consistent with the notion that social information systems such as Delicious is designed to facilitate the finding of general information, as tags created by other users increased the likelihood of finding these popular bookmarks (Hypothesis 2a, 3a). Therefore, experts got significantly more bookmarks in popular categories when using Delicious than using Google. In contrast, when using Google, experts utilized their domain expertise to generate “expert keywords” to search for information. As a result, they found more bookmarks that belong to the unique concepts (Hypothesis 3b). It is also possible that because Delicious is a general social system, it may not have indexed as much domain-specific information as

Google, which apparently had a much wider coverage of web sites that contain domain-specific information. However, it seemed that only experts were able to locate these unique bookmarks by coming up with keywords based on their domain expertise. For novices, the lack of domain expertise did not allow them to come up with as many keywords as experts, thus they were not able to find as many bookmarks in the unique categories as experts. Supported by the high number of bookmarks in popular categories selected by novices, Delicious did seem to help novices to find general information better than Google. This again demonstrated that social tags in Delicious did facilitate sharing of general information, even for novices who lack the domain expertise. However, finding unique information still needs domain expertise to facilitate the information search, as evidenced by the much higher number of bookmarks in the unique categories saved by experts than novices.

#### **5.2.4 Consensus on Tag Choices**

As the results above indicated, domain knowledge would influence how user search for information (3.2.2) and what information they collect from the Web (3.2.3). I further analyzed their tags in order to find out whether domain knowledge would influence their *interpretation* of information. Among the 48 participants, 3 participants have invalid tags (e.g., “bookmark 1”). The other 45 participants created 3046 tags in total. On average, every participant created 2.73 tags on each bookmark ( $SD = 1.76$ ). After getting rid of stop words and invalid tags, the number of distinct tags is 1384. As the number of distinctive tags was much fewer than the total number of tags, I speculated that the higher proportion of shared tags could be caused by: (1) social effect on tag choices in delicious, and/or (2) participants with similar knowledge background might have similar interpretation to information about one topic. In order to investigate which factor drove tag sharing, we performed a 2 (interface)  $\times$  2 (expertise) ANOVA using the number of users sharing each tag as dependent variable.

Results showed that the main effect of interface and expertise were significant ( $F(1,5528) = 54.75, p < 0.001$  and  $F(1,5528) = 7.65, p < 0.05$ ). The interaction effect of interface  $\times$  expertise was also significant ( $F(1, 5528) = 45.75, p < 0.001$ ). As shown in Figure 12, the interaction effect illustrated that experts using Delicious shared more tags

than novices ( $F(1, 2764) = 70.30, p < 0.001$ ), but no significant difference was found between experts and novices when they were using Google ( $p = 0.55$ ).

This result indicated that experts were more likely to agree with each other than novices in tag choices while they were in a social information system (Hypothesis 2b). Although it might seem surprising that experts had higher level of agreement on their tag choices even though they tended to search using their own queries, the result could be explained by their specific knowledge structures that influenced them to assign the same tags to the Web documents. This result could provide direct empirical evidence to the notion in the semantic imitation model that experts in the same domain were likely to share more common semantic representations of the same topic (Fu, et al., 2009). Therefore when experts were in a social environment, they tended to use similar tags as other experts. In contrast, novices tended to have more diverse interpretation to a topic, and might be more likely to use different tags to describe the bookmarks. In Google, experts and novices did not have this difference, possibly because of the mediation of the query suggestions provided by Google. Given that we did not collect data on query suggestions in this experiment, their effect could not be assessed; but their effects will be studied in future study.

### **5.2.5 Topic Learning Effect**

In order to investigate the learning effect brought by information search, I analyzed the keywords generated by participants before and after the search and the tags created by participants during search. Firstly, a 2 (interface)  $\times$  2 (domain expertise)  $\times$  2 (pre-/after-search) ANOVA was conducted to see whether there was a change in the number of keywords generated before and after search. The results showed that the difference between the pre-search session and the after-search session was significant ( $p < 0.10$ ), and the main effect of domain expertise was also significant ( $p = 0.05$ ). The main effect of interface and the interaction effects were not significant. Experts in general generated more keywords than novices, and all participants generated more keywords after search than before search. In order to see which group of participants had the biggest change after search, paired-sample t-test was conducted to compare the keywords generated by each participant in the four groups. The results showed that only in the

Expert-Delicious group, there was a significant increase in the number of keywords generated after the search ( $p < 0.10$ ) compared to the before search test (Table 4). The other three comparisons did not reach significance. This result partially suggested that experts' knowledge related with the concept of "financial crisis" was improved after searching using Delicious, but not for the other groups.

In order to check the knowledge consistency of the participants, I examined the semantic relatedness between each participant's tags, the keywords generated in pre-search tests, and the keywords generated in after-search tests using Latent Semantic Analysis (Laudauer & Dumais, 1997). The LSA calculations were performed through the web site at <http://lsa.colorado.edu>, using the general reading topic space with 300 factors. The document-to-document pairwise comparison was used to test the semantic relatedness between each set of keywords and tags. An example is shown in Table 5. The semantic distances between pre-search keywords and tags created in search were expected to indicate the social influence on participants' interpretations to the information content (which will affect the consistency in tag and keywords). A closer distance between pre-search keywords and tags would indicate less social influence brought by the search process. In other words, it is assumed that if participants relied more on their knowledge-in-the-head, the LSA score (indicating the semantic closeness) between their tags and topic-description keywords should be higher. In order to measure the knowledge change after search, I also analyzed the semantic consistency between pre-search keywords and after-search keywords. A larger semantic distance would indicate a bigger knowledge change, and vice versa.

2-way ANOVA was conducted to analyze the influence of domain expertise and search interfaces on knowledge consistency, using the LSA score between pre-search keywords and tags as dependent variables. In order to see the relationship between tags and topics induced from search, I also carried out a 2-way ANOVA using the LSA score between tags and after-search keywords as dependent variables. Another ANOVA was conducted to see the influence of domain expertise and search interfaces on knowledge change using the LSA score between pre-search keywords and after-search keywords as dependent variables.

As shown in Figure 13, the main effect of domain expertise was significant ( $F(1,41)=6.64, p<0.05$ ), but no other effect was significant. The semantic relatedness between experts' tags and pre-search keywords were higher than novices in two interfaces, which validated our previous result in 5.2.2 and 5.2.3 that domain experts relied more on their own knowledge, and thus their use of tags were more consistent with their domain knowledge than novices. Therefore, no matter what interfaces they were using, their interpretations to the selected information content were still influenced more by their own knowledge. Although it seemed surprising that experts had higher LSA score in Delicious than Google, this difference was not significant in post-hot analysis.

Similarly, only the main effect of domain expertise was significant in the comparison between tags and after-search keywords ( $F(1,41)=3.29, p<0.10$ , Figure 14). Experts had significantly higher consistency in their tag choices and keyword choices after search, indicating that the topics extracted from tags by experts were used to understand and interpret the information content. The comparison between pre-search keywords and after-search keywords also showed significant main effect of domain expertise ( $F(1,41) = 5.33, p<0.05$ ), but no other effect was significant (Figure 15). Experts therefore were supported to have less knowledge change after the information search, and novices had more knowledge change because of the exploratory information search. This result partly answered our question about the learning effect brought by exploratory information search (Hypothesis 2c) that domain novices have more knowledge change after search. These three comparisons also validated the assumption in the semantic imitation model (Fu, et al., 2009) that experts were more consistent with their choice of words to describe topics in their domain of expertise.

In order to further validate the learning effect, I merged the keywords generated by domain experts before and after search as the "expert" concept pool, and compared each novice's keywords with the expert concept pool. 2 (interface)  $\times$  2 (pre-/after-search) ANOVA was conducted using the LSA scores between novices' keyword and expert concept pool as dependent variables. However, only the main effect of pre-/after-search was marginally significant ( $F(1,20)=2.75, p=0.11$ ), and none of the other effects was significant (Figure 16). Generally, keywords generated by novices were semantically

closer to experts' knowledge after-search than pre-search, but the difference was not significant at the 0.05 level. Interestingly, novices were found to have more knowledge gain (higher semantic relatedness with expert concept) in Delicious after search, although the difference was marginally significant ( $p=0.11$ ). Novices using Google showed much less knowledge gain after search, and the difference between pre-search and after-search was not significant ( $p=0.89$ ). This part of result provided supports to the hypothesis 3c that the social information system would facilitate better learning effect. The results seemed to indicate that this effect was stronger in Delicious than in Google. However, given that our measure of knowledge was relatively simple (keyword generation), the marginally significant difference was promising. It is expected that with a more sensitive knowledge measure, more power could be obtained for the statistical tests.

### **5.2.6 Interview Results**

When being asked about the perceived usefulness of the two interfaces, most participants agreed that either of the interfaces was helpful for them to find information, and a few participants reported different opinions. Two of the participants using Delicious suggested that the interface was difficult to use, and three participants reported dissatisfaction about the search result provided by Delicious. Many users of Delicious, on the other hand, expressed their specific preference to the “websites suggested by people”, “peer review”, and “more specific and current information” provided by the system. Participants' opinion on Google are more consistent because they are very used to the search interface, but two of the 24 participants still expressed dissatisfaction about the search result because they “cannot find specific articles”. The interview results further validated our empirical data supporting that Google is better for general public to find general knowledge, but may not able to provide specific information to the information seekers. The average satisfaction rating about the search results from the 48 participants was 4.28 on a 5 point-scale.

Corresponding to the learning effect in 5.2.5, most participants reported they gained some knowledge from the 1.5 hour-search process, except 3 participants reported they only learned “a little” or “nothing”. This interview result helps supplement our empirical result in semantic analysis showing that novices in general gained knowledge

from searching in both interfaces. Although we did not specifically test the semantic change in experts' keywords, the higher number of keywords used in topic-description task after search by experts in Delicious group could also indicate some level of learning effect on experts.

### **5.3 Summary**

By conducting laboratory experiment with domain experts and novices, I found that domain expertise did influence participants' information behavior and the information sharing and information interpretation. The interaction effect between domain expertise and search interfaces was also validated by the results. Specifically, experts found more general information than novices by better interpretation of social tags in the tagging system; and experts also found more domain-specific information by generating more of their own keywords. Experts were also found to rely more on their own knowledge while novices were influenced more by the social web, as evidenced by their different search strategies and lower level of knowledge change. The results showed that although the social web could provide assistance to information seekers to some extent, domain expertise is still important in guiding them to find and evaluate the information.

Referring back to the hypotheses, our answers to each of the hypotheses are:

Hypothesis 1. Domain experts saved more bookmarks and created more tags than novice in the social information system. Domain experts used more keyword search, while domain novices used more tag search.

Hypothesis 2. In the social information system, experts were able to find similar number of popular information as well as less popular information, but novices were not able to find less popular information. Domain experts had more similar tags with each other, and domain novices had more diverse tag choices. Novices had more knowledge change than experts because of the less consistency in their keywords used.

Hypothesis 3. The social influence exerted on experts was weaker than novices because of the less unique concepts found by experts in Delicious than in Google.

Delicious still has limitations in supporting specific information search, but it is able to help domain novices finding more general information. Delicious was found to support better learning effect for domain novices, but the evidence was only marginally significant.



## CHAPTER 6. CONCLUSIONS

The empirical results from our study support the notion that domain expertise (domain experts vs. novices) and search interfaces (traditional search engines vs. social tagging systems) have a dynamic influence on users' information search behavior. Experts in general saved more information (bookmarks and tags) than novices. When we defined search efficiency as the number of URL visits per bookmark saved, experts had higher efficiency in Delicious and novices had higher efficiency in Google. Moreover, each search interface seems to facilitate information search in different ways. Specifically, we found that experts found more general information using Delicious and more unique information using Google, which was supported by the analysis of bookmarks and bookmark categories saved by both groups of users. At the same time, novices were able to find similar number of general information with experts, but much less unique information than experts in both interfaces. The results supported the claim that social information systems can facilitate the sharing of useful information among novice users, and social tags do seem to have strong potential to augment exploratory search of general information, even for users who have little knowledge on the topic. However, our analysis of unique information showed a discrepancy with previous studies. Domain experts still performed better in finding unique information in both interfaces, and the assistance provided by the present social search systems could hardly help domain novices to find specific information related to a topic.

The results on search strategies showed that experts used more keyword-based queries than novices in Delicious, while novices used more tag-based queries. This suggested that experts seemed to be capitalizing on their knowledge-in-the-head when performing exploratory search, but novices had to rely more on knowledge-in-the-social-web for their search. The semantic analysis provided further support to this fact as the experts had higher consistency in tag and keyword choices, and had less knowledge change after search.

Lastly, all our results consistently showed that experts conducted better information search in the social environment, as evidenced by the higher number of bookmarks saved, and the higher number of the topic description terms that they generated. The semantic analysis provided evidence supporting the better learning effect of novices induced by social information systems. But probably due to our limited experiment time and non-standard concept pool about the topic, the result was only marginally significant. Further research is still needed to provide stronger evidence for the learning effect.

## CHAPTER 7. DISCUSSION AND IMPLICATION

This study examines how domain expertise interacts with search environments to influence exploratory search performance. Because of the differences in people's knowledge structures, experts were driven more by a "top-down" process in the information search experiment by relying more on their knowledge-in-the-head. While this "top-down" process could bring more use of their own terms in keyword generation, it also limited their use of social tags in a social tagging environment. Novices on the other hand were found to use the shared knowledge of social systems (e.g., tags) more than experts, which also facilitated better learning (based on the LSA measures). Experts in general acquired more information than novices in Delicious, as reflected by the more bookmarks, more tags and better search efficiency by the experts. In short, social search environment allows the exchange of users' knowledge in their head through the media of social web. It is found that domain experts performed better as a "knowledge-receiver" in the experiment, but we did not find direct evidence showing that novices learned from the knowledge embedded in the folksonomies of the social tagging system. Future research should focus on how to provide a better interface for domain expertise to naturally "flow" into the social information systems as experts interact with them. Another important implication of this study is to provide empirical support to the semantic imitation model mentioned in 2.3. As supported by the model, domain experts will have more converged tag choices than novices. Results in tag sharing directly unveiled that experts are more likely to agree with each other on their tag choices, even if we did not control their information content at all. The high semantic relatedness between tags and after-search keywords also implied that the topic extracted from tags were used to understand the information content.

One possible future study is to examine how experts and novices understand and utilize recommended information differently. We have already provided empirical support to explain how they give tags differently and how they search differently. However, another major question of the current social information systems – low re-use of personal tags – still needs investigation. Although we assumed that providing "high

quality” tags from experts might be able to help the novices find more specific information, we still do not know how novices may interpret this information provided by experts who have different knowledge schema. A lot of potential directions could be involved in the study, including cognitive psychology, communication, and so on. For instance, Rader (2010) conducted a study to see the background influence of information producer, consumer and imagined audience on how the participants made use of the hierarchical structure created by the producer. She found that information consumers could find information most easily in hierarchies created by information producers who imagine their intended audience to be someone similar to them, independent of whether the producer and consumer actually shared common ground.

This study also indicated some limitations and future directions on the design of social search systems. A possible limitation of social tagging systems from the perspective of supporting exploratory search is the long tail of specific or unique bookmarks. By analyzing the bookmarks and bookmark categories in general and unique groups, we expected to see Delicious encouraging information sharing as well as knowledge exchange. However, the results showed the deficiency of social tagging systems in supporting the discoveries of domain-specific information. As expert users are more likely to generate their own search terms, especially for the unique concepts (lower frequency bookmarks or categories), it is likely that they will provide more unique and specialized tags. Therefore search results returned would have a smaller possibility to be shown on the popular page of Delicious. As a result, some useful results may be harder to be discovered because of its lower popularity.

Inspired by the different user behaviors in the two independent search interfaces, the combined use of traditional and social search environments might also be an interesting future topic to investigate. Social search systems, as well as traditional search engine, were both found to bring information gain in different conditions. In our findings, experts have the potential to provide information cues in a social search environment, but it is not easy for novices to pick up those cues because of their different knowledge backgrounds. In addition, experts could find more general information in the social search environment, but traditional search engine was better at providing unique

information. Another possible direction is to develop better data-mining techniques to recognize users based on their different behavioral pattern. Identifying the knowledge backgrounds of users would help personalize information recommendation, and also help improve the overall quality of the information space by utilizing higher quality input.

## REFERENCES

- Allen, B. (1991). Topic Knowledge and Online Catalog Search Formulation. *The Library Quarterly*, 61(2), 188-213.
- Ames, M., & Naaman, M. (2007). *Why we tag: motivations for annotation in mobile and online media*. Paper presented at the SIGCHI conference on Human factors in computing systems 2007, San Jose, California, USA.
- Bhavnani, S. K. (2001). *Important Cognitive Components of Domain-Specific Search Knowledge*. Paper presented at the TREC'2001.
- Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind*. New York: Cambridge University Press.
- Cattuto, C., Loreto, V., & Pietronero, L. (2007). Semiotic dynamics and collaborative tagging. *Proceedings of National Academy of Sciences*. 104 (5), 1461-1464.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In L. C. J. Barkow, & J. Tooby (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- Cosmides, L., & Tooby, J. (1994). Origins of domain-specificity: The evolution of functional organization. In L. H. S. Gelman (Ed.), *Mapping the Mind: Domain-specificity in cognition and culture*. New York: Cambridge University Press.
- Duggan, G. B., & Payne, S. J. (2008). *Knowledge in the head and on the web: using topic expertise to aid search*. Paper presented at the SIGCHI conference on Human factors in computing systems 2008, Florence, Italy.
- Fu, W.-T. (2008). *The microstructures of social tagging: a rational model*. Paper presented at the the ACM 2008 conference on Computer supported cooperative work, San Diego, CA, USA.
- Fu, W.-T., Kannampallil, T. G., & Kang, R. (2009). *A Semantic Imitation Model of Social Tag Choices*. Paper presented at the 2009 IEEE SocialCom.
- Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *J. Inf. Sci.*, 32(2), 198-208.

- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. T. (2007). Topics in Semantic Representation. *Psychology Review*, 114(2), 211-244.
- Heymann, P., Koutrika, G., & Garcia-Molina, H. (2008). *Can social bookmarking improve web search?* Paper presented at the international conference on Web search and web data mining.
- Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on the Search Tactics of Novice and Experienced Searchers. *Journal of the American Society for Information Science*, 44(3), 161-174.
- Kammerer, Y., Nairn, R., Pirolli, P., & Chi, E. H. (2009). *Signpost from the masses: learning effects in an exploratory social tag search browser.* Paper presented at the SIGCHI conference on Human factors in computing systems 2009, Boston, MA, USA.
- Kang, R., Fu, W.-T., & Kannampallil, T. G. (2010). *Exploiting knowledge-in-the-head and knowledge-in-the-social-web: Effects of domain expertise on exploratory search in individual and social search environments.* Paper presented at the SIGCHI conference on Human factors in computing systems 2010, Atlanta, GA.
- Kang, R., Kannampallil, T. G., He, J., & Fu, W.-T. (2009). *Conformity out of Diversity: Dynamics of Information Needs and Social Influence of Tags in Exploratory Information Search.* . Paper presented at the HCI International, San Diego, CA, USA.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York: Cambridge University Press.
- Krause, B., Hotho, A., & Stumme, G. (2008). A Comparison of Social Bookmarking with Traditional Search. *Advances in Information Retrieval*, 101-113.
- Laudauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- MacInnis, D. J., & Jaworski, B. J. (1991). Enhancing and measuring consumers' motivation, opportunity, and ability to process brand information from ads. *Journal of Marketing*, 53, 1-23.

- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Commun. ACM*, 49(4), 41-46.
- Millen, D. R., Yang, M., Whittaker, S., & Feinberg, J. (2007, Sept 26-28, 2007.). *Social bookmarking and exploratory search*. Paper presented at the ECSCW 2007, Limerick, Ireland.
- Morrison, P. J. (2008). Tagging of and searching: Search retrieval effectiveness folksonomies on the World Wide Web. *Information Processing & Management*, 44(4), 1562-1579.
- Pirolli, P. (2004). The Infoclass Model: Conceptual Richness and Inter-Person Conceptual Consensus about Information Collections. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society*, 11, 197-213.
- Pirolli, P. (2009). *Information Foraging Theory: Adaptive Interaction with Information*. New York: Oxford University Press.
- Rader, E. (2010). The Effect of Audience Design on Labeling, Organizing, and Finding Shared Files. Paper presented at the SIGCHI conference on Human factors in computing systems 2010, Atlanta, GA, USA.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory and Cognition*, 21, 803-814.
- Rouet, J.-F. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers*, 15(3), 409-428
- Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). *The cost structure of sensemaking*. Paper presented at the INTERACT '93 and CHI '93 conference on Human factors in computing systems.
- Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., et al. (2006). *Tagging, communities, vocabulary, evolution*. Paper presented at the 2006 20th anniversary conference on Computer supported cooperative work.
- Thom-Santelli, J., Muller, M. J., & Millen, D. R. (2008). *Social tagging roles: Publishers, evangelists, leaders*. Paper presented at the SIGCHI conference on Human factors in computing systems 2008.



- White, R. W., Drucker, S. M., Marchionini, G., Hearst, M., & schraefel, m. c. (2007). *Exploratory search and HCI: designing and evaluating interfaces to support exploratory search interaction*. Paper presented at the CHI '07 extended abstracts on Human factors in computing systems, San Jose, CA, USA.
- White, R. W., Dumais, S. T., & Teevan, J. (2009). *Characterizing the influence of domain expertise on web search behavior*. Paper presented at the Second ACM International Conference on Web Search and Data Mining, Barcelona, Spain.
- Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, 37, 3-15.
- Wilson, T. D. (1997). Information behaviour: an interdisciplinary perspective. *Information Processing & Management*, 33(4), 551-572.
- Zhang, X. M., Anghelescu, H. G. B., & Yuan, X. J. (2005). Domain knowledge, search behaviour, and search effectiveness of engineering and science students: an exploratory study. *Information Research*, 10(2).

## TABLES AND FIGURES

Table 1. 2 × 2 experimental design

	<b>Google</b>	<b>Delicious</b>
Expert	12 participants	12 participants
Novice	12 participants	12 participants

Table 2. Statistics about knowledge tests

		<b>Self Report (5-point scale)</b>	<b>Knowledge questionnaire (20 items)</b>	<b>Specific questions (10 questions)</b>	<b>Average # of keywords Pre-search</b>	<b>Average # of keywords After-search</b>
Mean	Expert	3.9	10.83	5.08	21	26.54
	Novice	2.88	8.5	3.67	14.97	19.59
p-value		0	0.011	0.006	0.095	0.15

Table 3. Examples of the categories of bookmarks saved by the participants

	<b>Categories</b>	<b># of users</b>	<b># of bookmarks</b>
Popular	Cause	11	154
	Effects	10	61
	History	7	35
	How to deal/end/to do/reaction	7	19
Unique	AIG	2	15
	Wall street	2	12
	CMBS/CDO	1	6
	Big three/ US auto	1	2

Table 4. Comparison between pre-search and after search topic-description tests

	<b>Number of keywords generated</b>		<b>Standard error</b>	<b>p-value</b>
	<b>Before-search</b>	<b>After-search</b>		
Expert-Delicious	18.91	22.58	2.35	0.095
Expert-Google	23.08	30.5	8.55	0.404
Novice-Delicious	12.75	18.08	3.56	0.162
Novice-Google	15.75	19.33	4.63	0.456

Table 5. An sample set of pre- and after-search keywords and tags

<b>Pre-search keywords in topic-description task</b>	current recession GMAC Freddie Mac Fannie Mae Government subsidization Subprime mortgage lending government errors low interest rates problems in auto industry defaulting on loans subsidization (governmental) recessions cutting back market downturn economic problems
<b>Tags created to all bookmarks in experiment</b>	yahoo wikipedia wall street visualization unemployment swaps subprime stocks stearns steagall solutions simple region recovery recession rates quotes quote politics online news myths mortgage money market mac list lehman krugman jones investments investing impact housing history graphics graphic google glass freddie financial finance
<b>After-search keywords in topic-description task</b>	bankruptcy inflation Lehman Brothers AIG bonuses subprime mortgage crisis Auto bailout recession depression deficit federal reserve bank Freddie Mac Fannie Mae Bailout Plan Investment Banks Bubble bursting Unstable financial market housing market decline Alan Greenspan predictions liquidation

Figure 1. The semantic imitation model: In the figure, existing tags (T1, T2, T3) act as cues for related topics (C1, C2, C3) in the topic inference process, and later lead to extraction of gist concepts Cw, Cx, Cy, and Cz.

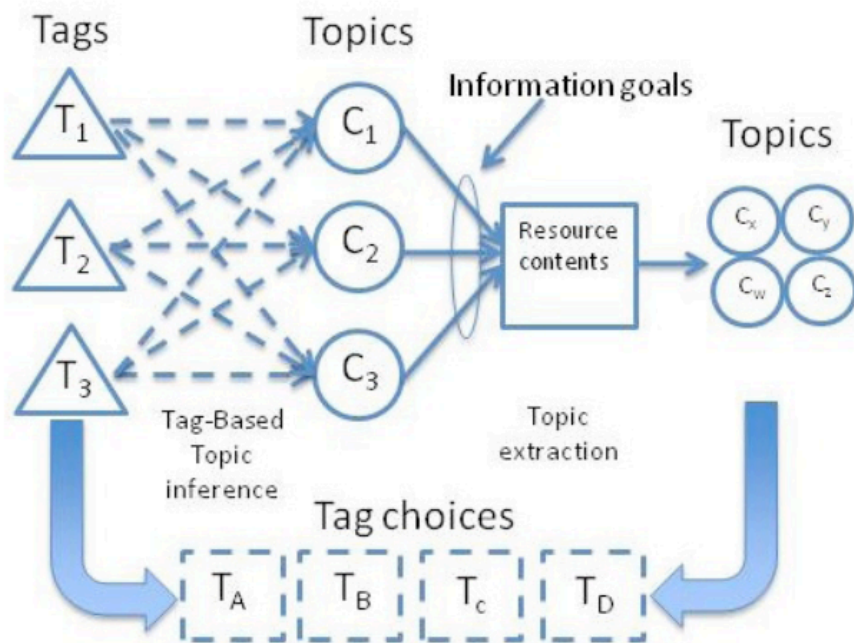


Figure 2. Top: Scatter-plots of tag proportions against tag choice cycles by the semantic imitation model when there were simulated domain experts (left) and novices (right).

Bottom: Scatter-plots of choice proportions of each tag assigned to a single-topic document simulated by the semantic imitation model when there was simulated domain experts (left) and novices (right).

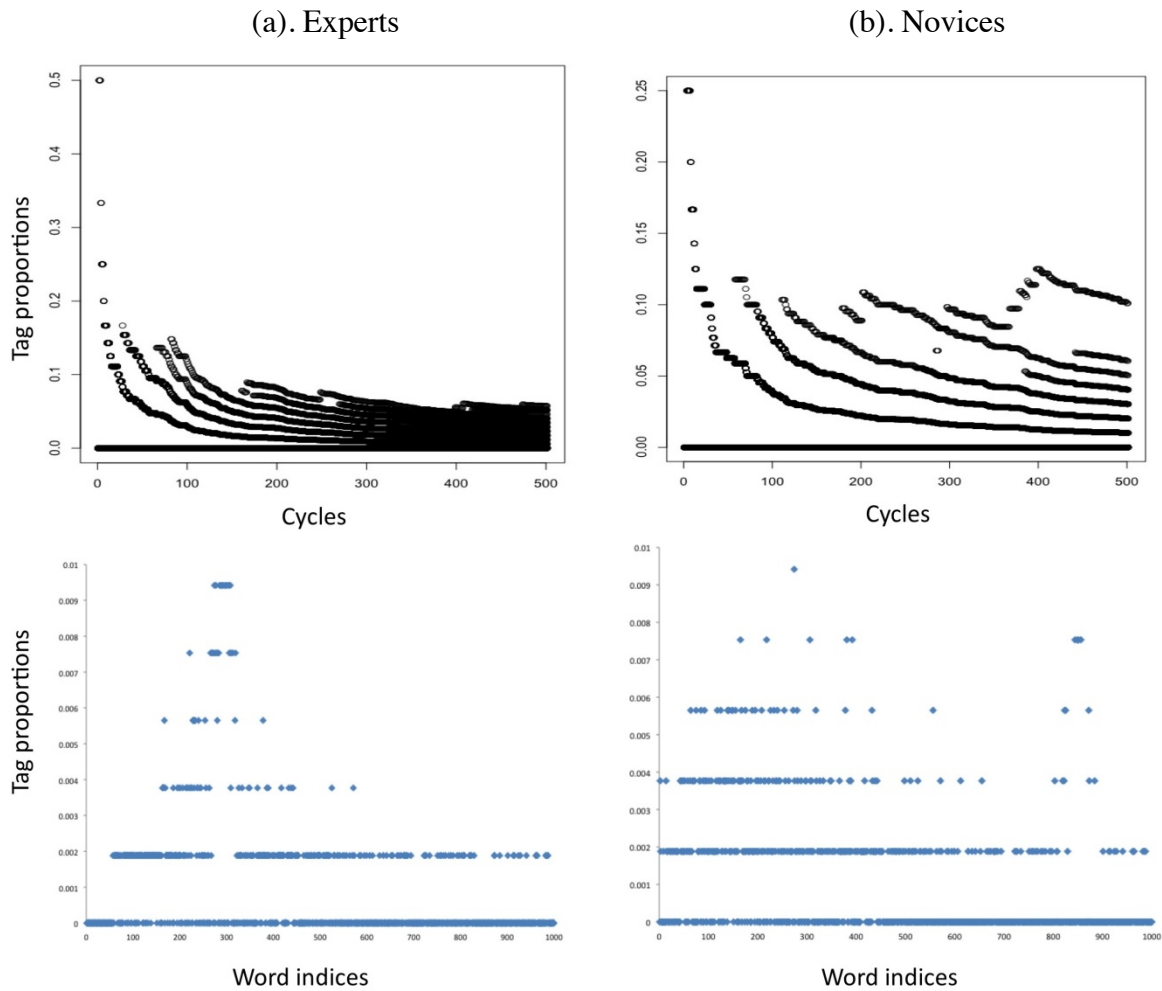


Figure 3. Mean number of unique tags assigned to each book by participants in the social and nominal group

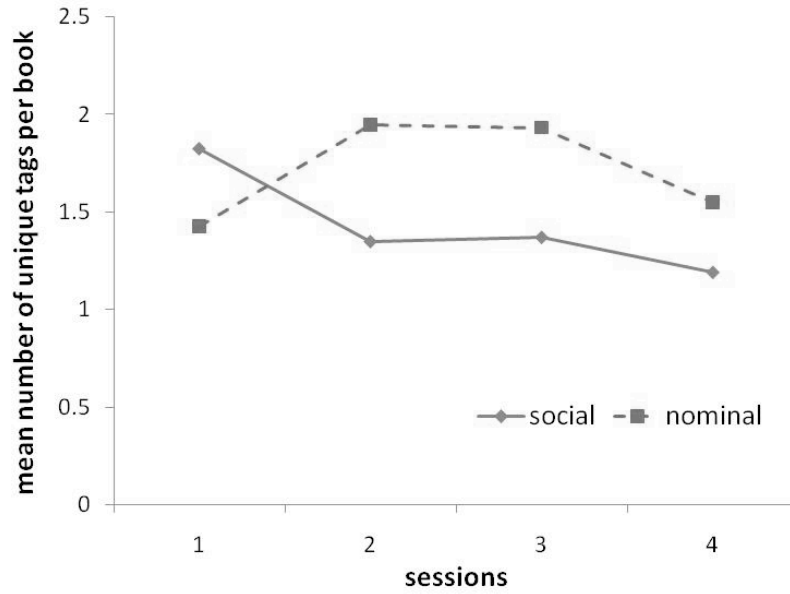


Figure 4. LSA scores for new tags created across sessions in the social and nominal groups, broken down by whether the books were selected under the same or different search tasks

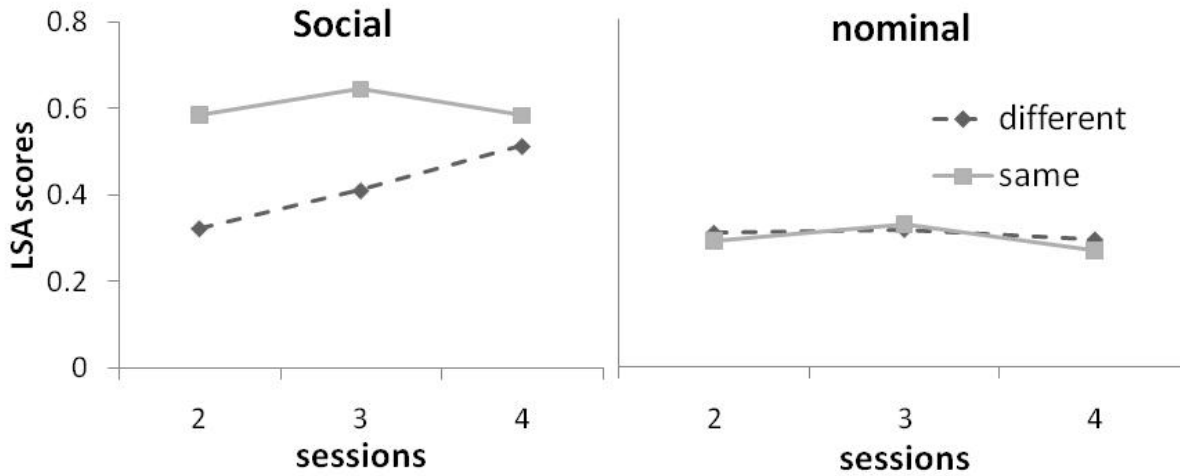


Figure 5. Search strategies in Google and Deliciou: In the top figure, users can enter keywords to the search box, then Google will return a list of matched results; in the bottom figure, users can either enter keywords in the keyword box, or click on any of the tags to search, then Delicious will return a list of URLs together with title, and user generated abstract and tags.



Figure 6. The experiment procedure

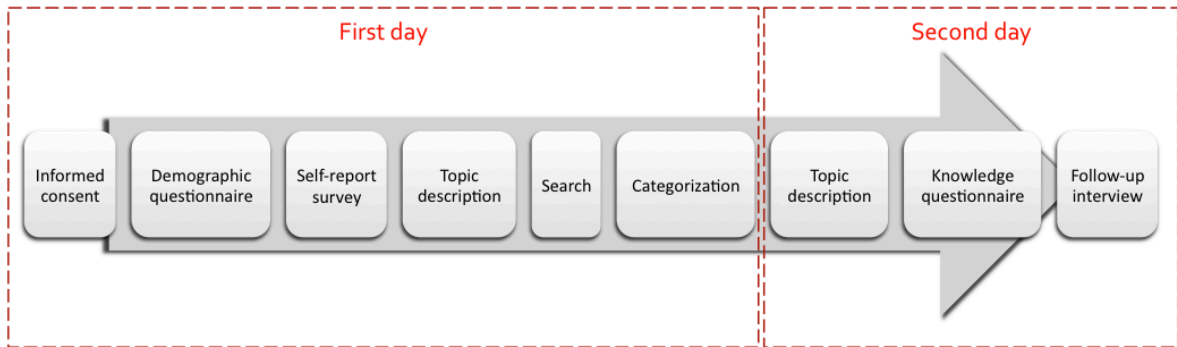


Figure 7. Comparison of search behavior between experts and novices in two interfaces

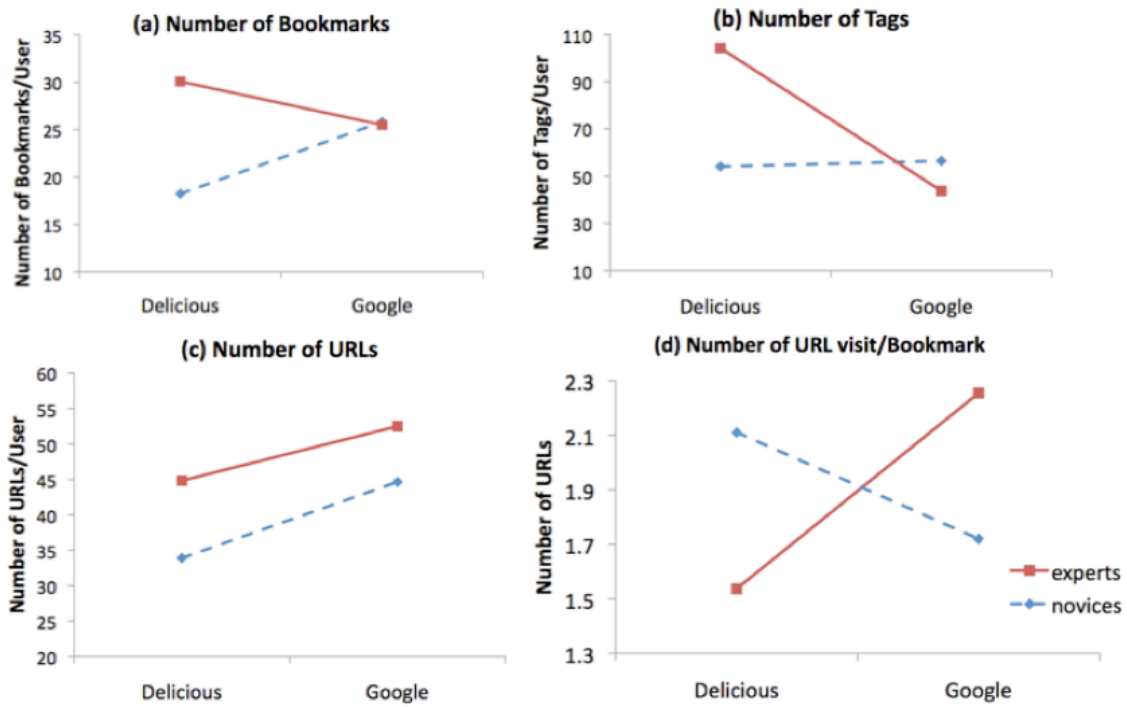




Figure 8. Search strategies employed by four groups of participants

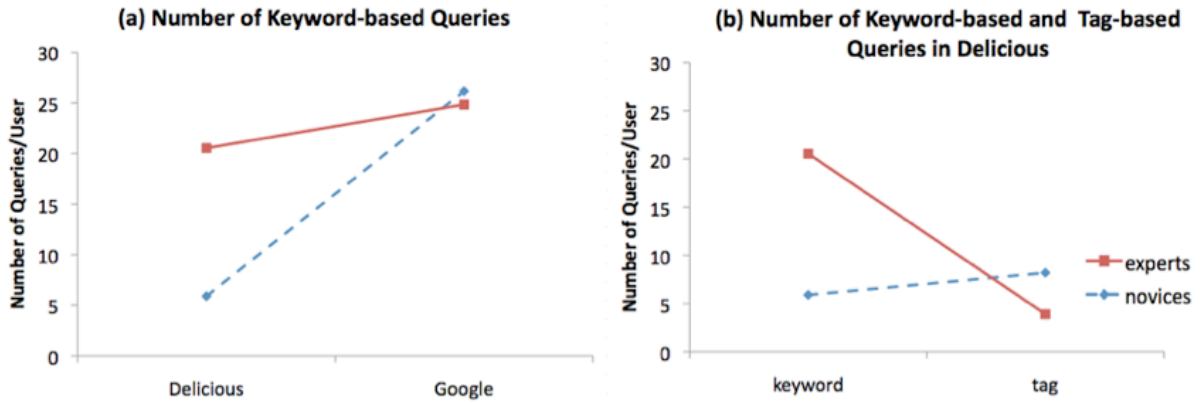


Figure 9. The frequency-rank plot of bookmarks saved by all participants

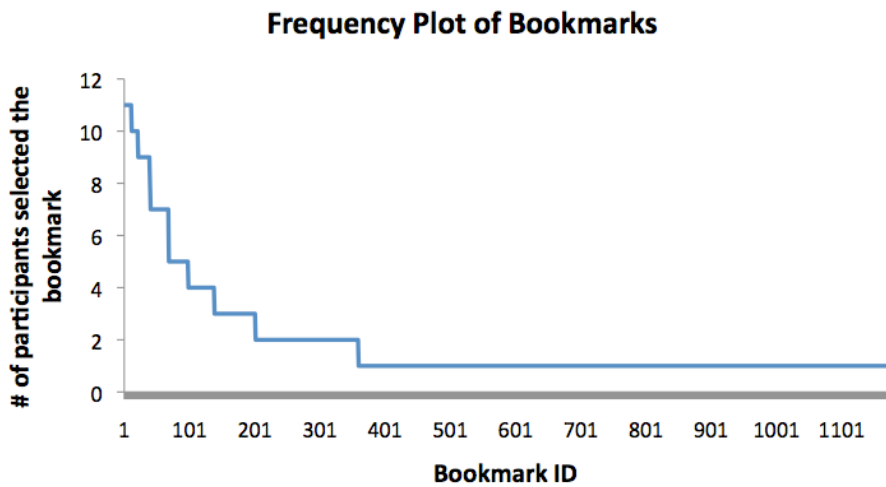


Figure 10. The impact of expertise and interface on bookmark sharing

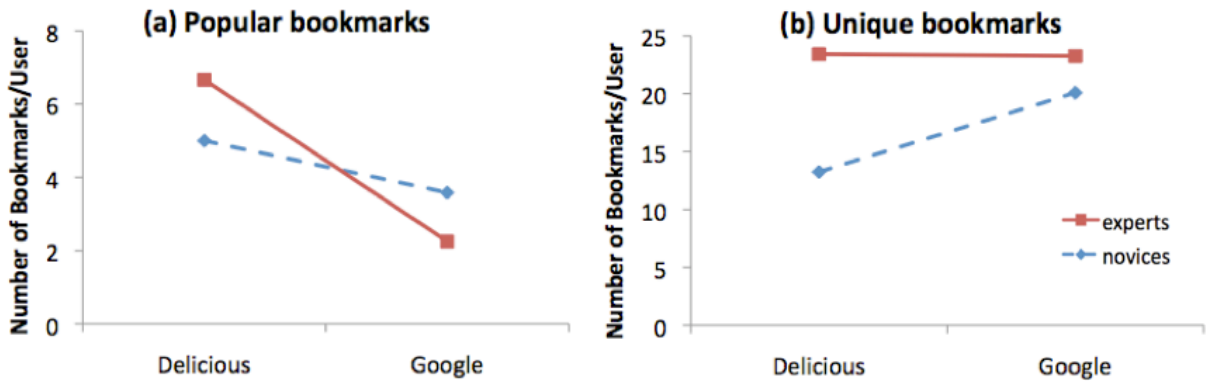


Figure 11. The impact of expertise and interface on bookmark sharing

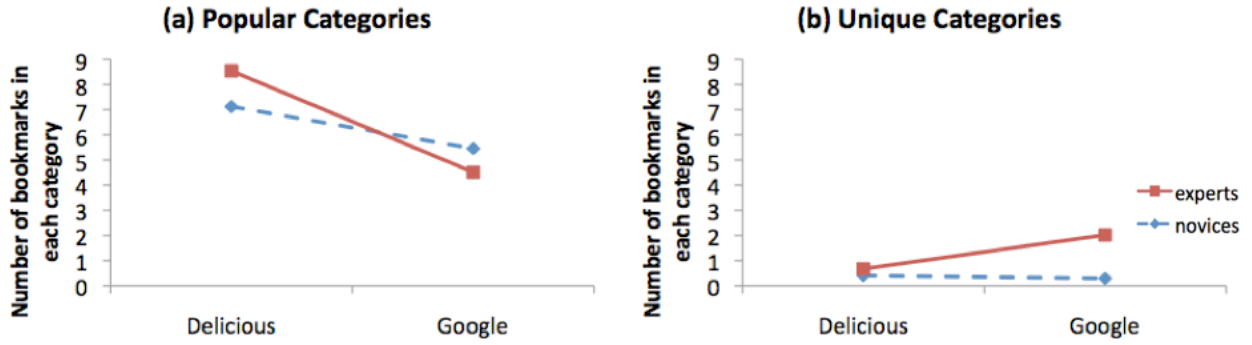


Figure 12. Tag sharing in Delicious and Google

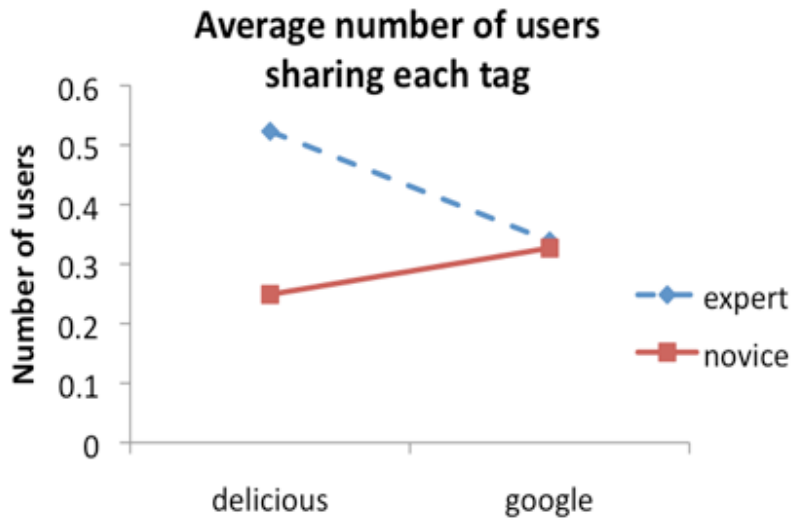


Figure 13. The average semantic relatedness between pre-search keywords and tags for each participant

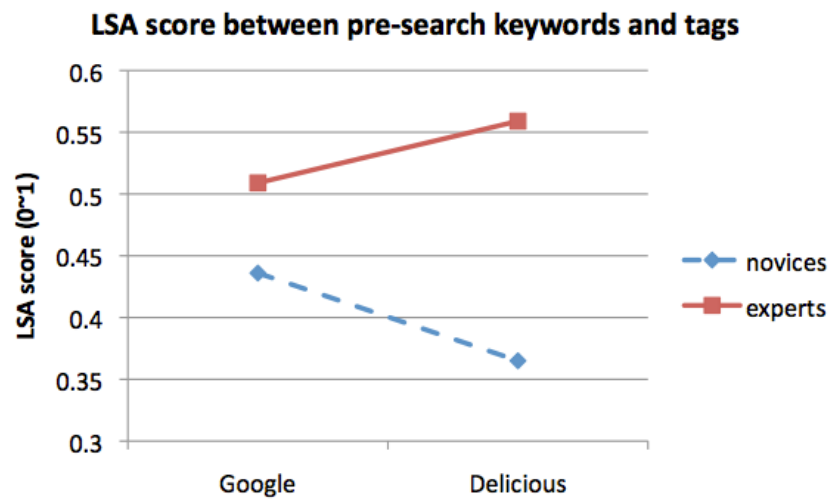


Figure 14. The semantic relatedness between tags and after-search keywords

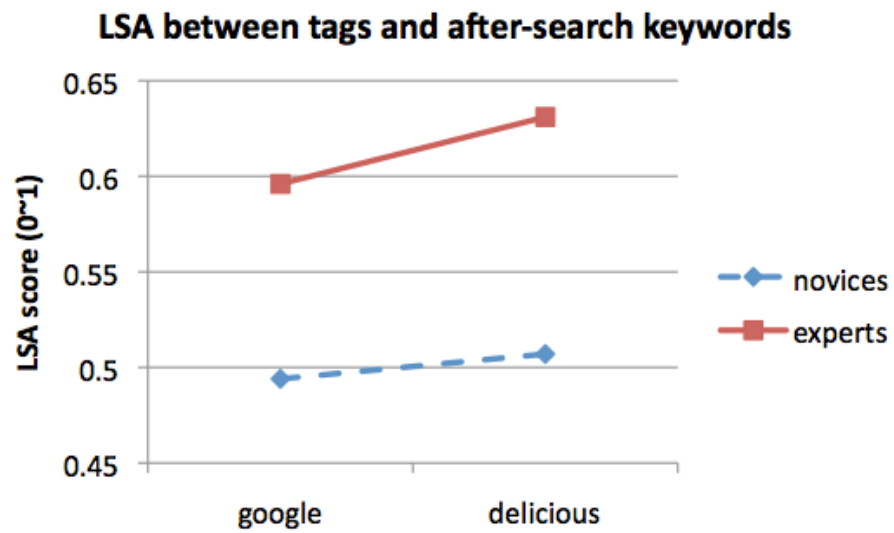


Figure 15. The average semantic relatedness between pre-search and after-search keywords

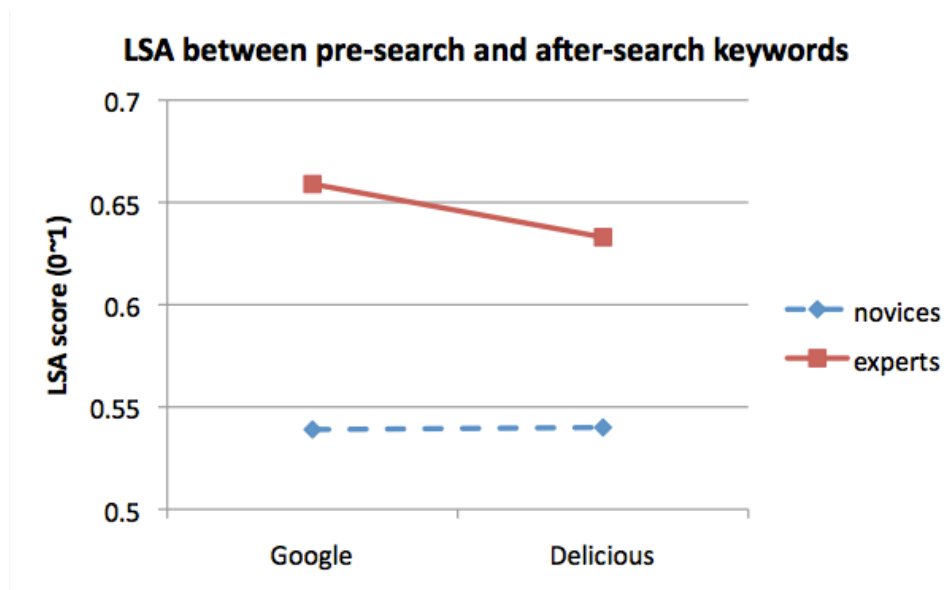


Figure 16. The semantic relatedness between novices' keywords and experts' concepts before and after the information search task

