Supporting Exploratory Information Seeking

by

Mona Haraty
Bachelor of Science, University of Tehran, 2008

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
School
of
Interactive Arts and Technology

© Mona Haraty 2010

SIMON FRASER UNIVERSITY

Fall 2010

APPROVAL

Name:	**Monasadat Haraty**
Degree:	**Master of Science**
Title of Thesis:	**Supporting Exploratory Information Seeking**

Examining Committee:

Chair: _____

Dr. Marek Hatala

Associate Professor of School of Interactive Arts and Technology

_____

Dr. Brian Fisher
Senior Supervisor
Associate Professor of School of Interactive Arts and Technology

_____

Dr. John Dill
Supervisor
Professor Emeritus of School of Interactive Arts and Technology

_____

Steve DiPaola
External Examiner
Associate Professor of School of Interactive Arts and Technology

Date Defended/Approved: _September 3, 2010_

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Simon Fraser University
Burnaby, BC, Canada

Last update: Spring 2010

ABSTRACT

Exploratory search is a sensemaking activity that involves information seeking and iterative development of mental model of the domain under exploration. It often begins with a vague and evolving information need that is multidimensional.

We designed and developed a web-browser extension to facilitate exploratory web search aiming at transforming the search activity into a meaningful learning activity. The design is based on the proposed multi-threaded model of exploratory search. According to this model, exploratory search is a multi-threaded process as the user has multiple concurrent sub-goals addressing different aspects of her information need.

A case study is conducted to evaluate the design and investigate how the proposed model can provide support for sensemaking activities involved in the exploratory web search.

Keywords: exploratory search, sensemaking, Information seeking, reflection, cognitive architecture, externalization, cognitive load.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1: INTRODUCTION

The overall purpose of this thesis is to improve our understanding of sensemaking processes in exploratory web search as well as designing and building a system that can support exploratory web search by facilitating sensemaking of the exploration process as well as facilitating sensemaking of the information.

We suggest a multi-threaded model for exploratory search behavior in which the user has multiple thought streams following multiple concurrent goals and sub-goals, which are emerging during the exploration. We propose the concept of *"Exploration Thread"* as a representation of a user's thought stream around one aspect of his information need. According to the sensemaking models' terminology, we can define each thread as a representation for organizing the information related to the exploratory search process for reducing the cost of making sense of the information.

We are using visualization as a means for externalizing the Exploration Threads. Each Exploration Thread is a visualization of a part of the user's trail that has a tree structure in which the nodes are representing the web-pages and a link from node 'A' to node 'B' represents the user's path from 'A' to 'B'. The user can impose structure on the visualization. The structures imposed by the user can represent how the user is thinking about the path that she is taking and how she relates web pages to each other according to her goals and sub-goals

for each search task. In this approach, the construction of representation for organizing the information is integrated into the exploration process and hence it doesn't impose too much extra effort on the user.

1.1 Background

Making sense of a large body of unorganized information is the primary task of analysts. Sensemaking happens when people encounter new problems and situations in which their current knowledge is insufficient for addressing them. It involves finding important structures among a seemingly unstructured environment to serve the task at hand. The complexity of a sensemaking task can be increased due to several factors including lacking knowledge of a domain and unclear information need. Lacking knowledge of a domain is experienced in everyday web search activities, when people seek to understand something outside the area of their expertise. As a result, they are spending a considerable amount of time for gaining the appropriate knowledge in order to understand the answer to their information needs. Examples of such activities are buying a device such as a new laptop or camera. Another example is looking up a new word in Wikipedia, especially when the word is completely out of the user's domain knowledge which makes the user look up several other words to be able to understand the first word.

The second factor in increasing the complexity of a sensemaking task is unclear information need, which is the main characteristic of exploratory search. In designing and building systems aimed at facilitating sensemaking, one should consider reducing the complexity of the sensemaking task created by these

factors. Facilitating the development of a mental model of the domain can compensate for the inefficiencies caused by the lack of domain knowledge. Moreover, making the user aware of the changes such as evolution of his interests over time can help in structuring the information need, and hence compensates for the inefficiencies caused by the unclear information need. However, the current web browsers do not provide any specific support for users who hardly have well defined questions to get answer for or even any particular goal in mind. Having no particular goal or defined question simply happens when the user is involved in an "open-ended browsing" according to Shneiderman (Shneiderman, 1997). In an open-ended browsing or as Marchionini calls it exploratory search (Marchionini, 2006), there is an implicit goal of gaining an understanding of a general subject area that has not been supported well in the existing tools.

Exploratory search is an example of sensemaking process in which determining goals of a task is part of the sensemaking process. Every sensemaking task has a corresponding target task, which determines the goal of the sensemaking. "Finding the search goal" or "clarifying the question/information need" are two possible target tasks in exploratory search. Exploratory search can also be explained as one of the tasks in the large class of problem solving tasks (A. Newell, 1994). Indeed, it is an ill-structured problem that requires additional sources of knowledge to better understand the starting state, to define the goal state, and to understand what actions can be taken at any given state of the exploration (Simon, 1973). The problem solving perspective on exploratory

search can also inform the design of systems aimed at supporting sensemaking in exploratory search. For example, not having a well-defined goal state implies that the design should consider and facilitate the process of clarifying and managing goals for the user.

Conducting research in its early stage is an example of exploratory search (Pedersen, Gyllstrom, Gu, & P. J. Hong, 2010). Some of the properties by which early stage of research can be characterized are (Pedersen et al., 2010):

1. The research question and the information need are still in formation,

2. The process is fragmented,

3. The exploration theme is changing slightly as the user learns more about the domain,

4. Premature structures are applied on the pages, although they are the result of an outdated understanding as the user is learning and expanding his knowledge of the domain.

In early stage research, as well as other sensemaking domains such as intelligence analysis, the insight comes from the process of exploration rather than just from the end result. In some cases, the exploration process sets the stage for reaching to and understanding the end result and in other cases where the user is not looking for any end result, the process is the only source of insight.

In exploratory web search, facts related to a specific search topic are distributed across several web pages. Likewise, the required information for

understanding a domain is distributed in papers, books, and web pages. Therefore, it is challenging for a new researcher in an interdisciplinary domain to get a sense of the domain considering the distribution of facts across different disciplines outside his area of expertise. Effects of domain knowledge on search tactic and its evolution have been studied for years (Wildemuth, 2004; White, Dumais, & Teevan, 2009). Goal sequencing strategies and important URLs are considered as the domain specific knowledge that affects the effectiveness and efficiency of the search process.

This thesis reports design and evaluation of Explore-It, which is a Firefox extension that has been designed and implemented for facilitating exploratory web search through supporting reflective sensemaking based on a multi-threaded model of exploratory search.

Explore-It consists of two components including an Exploration Path and a tag cloud of user's interest that collectively address the goals discussed above. Exploration Path is a visualization of the path explored by the user; it is created in a mixed initiative way to represent the ongoing thought streams and to represent the implicit sub-goals that have been followed in the exploration. The second component is a tag cloud, which contains the keywords used in the queries, threads' names and tags, the user's comment, and web page's title. Tag cloud represents the user's interest and its evolution during the exploration to make the user aware of the changes in his interests that can help in managing the information need and goals in exploration.

1.2 Thesis Contribution

The main contribution of the thesis is a multi-threaded model for exploratory web search, which proposes that users follow multiple parallel thought streams in their mind while being involved in exploratory search and they have multiple concurrent goals to achieve in an exploratory search. This model can improve the current cognitive architectures such as ACT-R to provide support for exploratory search. Other contributions of this thesis is summarized in the following:

- Building on the existing sensemaking models and extending them to consider sensemaking of the process rather than just focusing on sensemaking of the data

- Mediating reflection-in-action by having the user think about the visited web pages while searching

- Incorporating a novel mixed initiative approach in constructing the sensemaking representation by integrating the sensemaking and reflection activities into the activities related to searching

- The result of this study will improve the design of exploratory search supporting tools for providing better support for the sensemaking activities in the exploratory search

1.3 Thesis Outline

Chapter 2 begins by providing some background on exploratory information seeking and it will continue by covering the sensemaking models, cognitive

architectures, theories in learning, and reflection in/on action, all of which worth

to be considered in designing an exploratory supporting tool with the goal of

facilitating sensemaking. Chapter 3 reviews the related works including existing

tools and systems with similar goals in general. Chapter 4 introduces the design

guidelines suggested by distributed cognition theory and the principles of

reflection in/on action. Chapter 5 presents the design and implementation of the

system as well as the rationale behind all the design decisions that mostly comes

from the guidelines discussed in chapter 4. Chapter 6 focuses on the research

methodology by providing some background on the chosen methodology and

justifying the choice of method. The reports on the results of the case study are

presented and discussed in chapter 6. Finally, Chapter 7 summarizes the thesis'

contribution and discusses the future work.

2: THEORETICAL FOUNDATION

In this chapter we review concepts, models and other theoretical foundations required for understanding the design space of exploratory search support systems.

2.1 Exploratory Information Seeking

Exploratory search usually starts with an ill-defined query that does not necessarily represent the information needs of the searcher, as most of the searchers usually do not pursue any explicit goal. However, there is definitely an implicit goal of learning and making sense of a topic or a knowledge domain, which keeps the searchers continuing their explorations. Therefore, exploratory information seeking can be considered as an effective learning activity, that the lack of support tools can transform it into a frustrating activity. For example, information seeking in an interdisciplinary research area is a challenge that new researchers are constantly facing. This can be improved through facilitating the sensemaking process. Sensemaking is the process of searching for a representation that organizes information to reduce the cost of an operation in an information task (Russell, Stefik, Peter Pirolli, & Stuart K. Card, 1993).

Exploratory search is a type of search in which the information need is unclear for the searcher and it is different from fact-finding or question answering search tasks. Exploratory search is accompanied with sensemaking processes

including making sense of the information encountered and making sense of the process of exploration.

2.2 Sense-making Models

Sensemaking occurs when people encounter new and unfamiliar situations in which their current knowledge is insufficient for dealing with them. According to Russell et al., "sensemaking is the process of searching for a representation and encoding data in that representation to answer task specific questions"(Russell et al., 1993). It involves finding some structures in seemingly unstructured situation by developing successively more sophisticated representations and fitting information into representations in service of a task. Depending on the type of the task at hand, the representations are chosen or constructed in a way to reduce the cost of operations in an information task. In general, sensemaking involves the following steps: 1. Recognizing a knowledge gap; 2. Generating an initial structure of the knowledge needed to complete the task—concepts, relationships, and hypotheses; 3. Searching for information; 4. Analyzing and synthesizing information to create an understanding; 5. Creating a task product based on this understanding in the form of a report, decision, or problem solution (P. Pirolli & S. Card, 2005). As sensemaking involves searching for or creation of structured representations, systems aimed at supporting sensemaking should provide support for building those structures. Therefore, building such systems requires sensemaking models that provide enough detail of the sensemaking process.

Several sensemaking models for capturing the sensemaking processes have been proposed. Russell et al. recognized four main processes of sensemaking (Figure 1) (Russell et al., 1993).
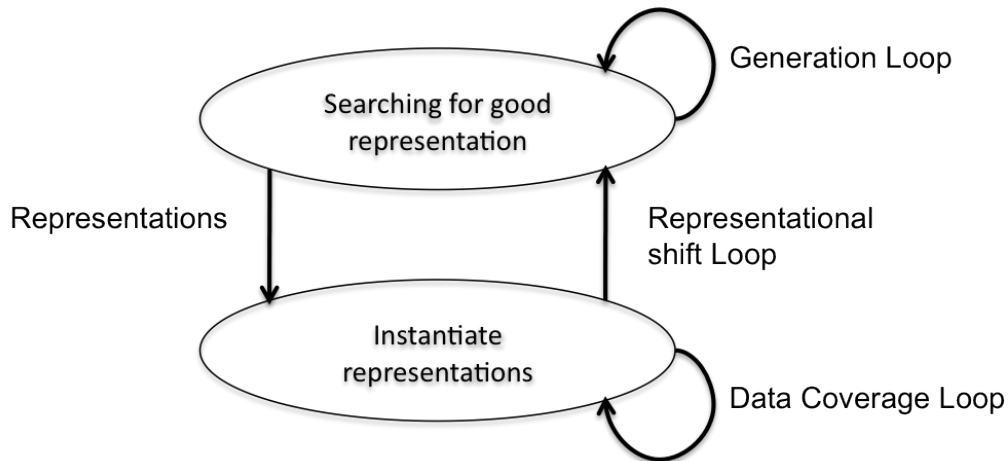
Figure 1: Learning Loop Complex in sensemaking tasks

1. Search for representation (structure): The sense-maker searches for and creates representations (structures or schemas) that can be used to organize the information needed for the task (generation loop).

2. Create instances of representations: The sense-maker identifies information of interest and encodes it in the representation (data coverage loop).

3. Modify representation: The sense-maker modifies the representations when the data cannot fit into the previously established representation (representation shift loop).

4. Consume instantiated representations: The sense-maker consumes the instantiated representation and uses it in performing the task.

According to Russell et al.'s sensemaking model, the sensemaking representations are constructed and changed in order to reduce the cost of operations in an information task.

Pirolli and Card derived a "notional" model of sensemaking through a cognitive task analysis (P. Pirolli & S. Card, 2005). According to this model, the overall sensemaking process consists of information gathering, representation of the information in a schema that aids analysis, developing insight through the manipulation of the representation, and creating some knowledge product or direct action. The model further separates two loops of activities:

- An information foraging loop that involves searching for information, filtering it, and reading and extracting information into some schema;

- A sensemaking loop that involves iterative development of a mental model from the schema that best fits the evidence.

Below are two common examples of sensemaking tasks:

Example1: "A business analyst gathers, analyzes, and uses information about a product and its alternatives, customers, and competitors, and makes planning suggestions for the organization. The task is familiar to the analyst, but the domain and the product might be new from time to time. The analysts need to extract related concepts and entities from the information they found and create an understanding of various relationships and make reasonable suggestion based on the identified patterns and relationships."

Example2: "A patient has been diagnosed as having high blood pressure. He needs to learn about the condition. The problem and the domain are both new to this patient. He needs to find out about the causes, consequences, treatment options, influences, and so on. The major task for him is to develop a mental model of the condition and relates it to his previous knowledge structure so that he could make decisions and take actions"

The two examples are adopted from (P. Zhang, 2010)

2.3 Cognitive Architectures

In order to support sensemaking, construction of sensemaking representations should be facilitated for the information seeker during the exploration process. According to distributed cognition theory (Hollan, Hutchins, & David Kirsh, 2000), people tend to offload their cognitive burdens in cognitively demanding tasks. Exploratory search is a cognitively demanding task due to its complexity as the user is following multiple thought streams in parallel.

Therefore, one of the goals of the model that we propose for exploratory search is to provide support for externalizing the internally constructed sensemaking representations. From the distributed cognition perspective, this goal can be interpreted as providing support for offloading the cognitive load. In order to facilitate the externalization of the sensemaking representations, we have to understand the cognitive processes that are involved in the exploration process. Cognitive architectures, which are theories for understanding and simulating human cognition, can help us in understanding the cognitive processes in the exploratory search.

ACT-R (Adaptive Control of Thought--Rational) is a cognitive architecture, which focuses on understanding how people organize knowledge and produce intelligent behavior. ACT-R has a distinct goal-memory in addition to the declarative and procedural memory, which is a hierarchy of user's intentions in cognitive activities such as information foraging. A threaded cognition model within the ACT-R framework that describes multi-tasking behavior has been proposed by (Salvucci & Taatgen, 2008). Considering the exploratory search as a multi-tasking behavior, tasks are the search activities aimed at fulfilling various aspects of the user's information need.

2.4 Learning Theories

Research in cognition and learning theory can also provide important insight for understanding sensemaking. Learning activities are similar to the sensemaking activities in terms of cognitive processes that the learner/sense maker undertakes. Three of the learning theories that are closely related to

sensemaking are Assimilation Theory (Ausubel, Novak, & Hanesian, 1978), Schema Theory (R. C. Anderson, 1984), and Generative Learning Theory (WITTROCK, 2010).

According to Assimilation Theory, meaningful learning is a learner-controlled process in which a new piece of information is assimilated to an existing relevant aspect of the learner's knowledge structure. Knowledge can be considered as a schema stored in human memory with interconnected concepts and relationships, organized in a meaningful way (R. C. Anderson, 1984; Rumelhart & Ortony, 1977). The learner is actively involved in placing new pieces of information in the existing schemas or constructing new schemas and modifying the existing ones in case the new pieces of information do not fit into the existing schemas. The construction of knowledge structures is the key to sensemaking as well. Assimilation Theory makes the following claims:

1. The development of new meanings is built on prior knowledge,

2. The learner's cognitive structure is organized hierarchically with more general concepts occupying higher levels in the hierarchy,

3. Meaningful learning clarifies the relationships between concepts.

The procedures that a learner goes through to learn something is essentially the same as the procedures that a sense-maker goes through to make sense of concepts and relationships between them. Therefore, meaningful learning can be considered as a form of sensemaking. If we apply the Assimilation Theory's claims to sensemaking, we can conclude that new meanings are derived based

on the prior knowledge about the task situation. However, the question of how the cognitive structure is organized in a sensemaking task cannot be easily answered from the claims of Assimilation Theory. The knowledge structure or the structure of sensemaking representation is not necessarily arranged hierarchically. The structures highly depend on the nature and complexity of the task and the subject's domain. The major concern of this thesis is to understand the structure of these representations for exploratory web search.

The Schema Theory states that knowledge is stored in human memory as schemas with interconnected concepts and relationships (R. C. Anderson, 1984; Rumelhart & Ortony, 1977). Some of the principles of Schema Theory are as follows:

1. Schemas change as new information is acquired

2. Prior knowledge is necessary for new knowledge

3. Each individual's schema is unique and dependent on that individual's experiences and cognitive processes

4. Learners feel internal conflict whenever new information does not fit into the previous schemas

The difference between Assimilation Theory and Schema Theory is in the organization of knowledge structure. The Assimilation Theory assumes a hierarchical structure for organization of knowledge as opposed to the networks of propositions in the Schema Theory. According to Schema Theory, in order for

learners to learn and remember certain knowledge, it has to be organized mentally as an interconnected network.

Generative Learning Theory emphasizes that the learner is involved in constructing two types of meaningful relationships; one between information and experience and the other among the parts of information (Grabowski, 2004). Generative Learning Theory identified two types of generative learning activities:

- Creating organizational relationships between different elements in the environment including creating titles, headings, summaries, graphs, and tables

- Creating relationships between the external representation or stimuli and the memory components or internal representations, including metaphors, analogies, and examples

These generative learning activities represent sensemaking activities such as extracting concepts from the new information and positioning them in the existing structures.

In sum, structural knowledge is one of the central concepts in learning and sensemaking. Structural knowledge plays an important role in understanding information and modifying the existing knowledge. Therefore, drawing the learner's attention to structural aspects of the information can improve acquisition of knowledge. Visualizing information and the connections between pieces of information helps the acquisition of structures; therefore, utilizing visualization in providing support for sensemaking is one the primary design decisions of

sensemaking support systems. Moreover, learning theories' insights into "how people learn new pieces of information" can be used for supporting sensemaking in exploratory search, where searchers are constantly facing new pieces of information.

2.5 Reflection in/on search process

2.5.1 Reflection-in-Action (mediating reflection)

Reflection is also known as metacognitive thinking (Brown, Bransford, Ferrara, & Campione, 1983). Metacongnition is an important component for success in learning new domains, solving problems, effectively utilizing prior knowledge, and organizing information and resources (Bransford & Stein, 1993).

According to Schön, a practitioner's reflection-in-action is bounded by the "action-present", that is "a zone of time in which action can still make a difference to the situation". The action-present depends on the characteristics of the action, including the pace of the activity and the situational boundaries. In order to determine this time frame for exploratory search, first, the unit of action in which the user can reflect should be determined. We can specify several units of action in the exploratory search process that the user can reflect on them.

3: RELATED WORK

Several research communities have either addressed problems similar to ours with different focuses and various perspectives or proposed solutions similar to ours to different problems. One of the similar problems is "getting lost in hyperspace" while browsing the World Wide Web, which leads to facing difficulty in re-finding or re-visiting web pages that have already been visited. Several researchers have used visualizations of browsing history as part of their solutions to the orientation problems in hyperspace and re-finding and re-visiting problems (Hightower, Ring, Helfman, Bederson, & Hollan, 1998; Wexelblat & Maes, 1999; Tabard, Mackay, Roussel, & Letondal, 2007; Brian Fisher et al., 1997; Ayers & Stasko, 1995; Doemel, 1995; Frécon & Smith, 1998). The common approach among all of them is visualizing the user's trails using tree and graph data structures (Turetken & Sharda, 2007). However, they are using different methods for expanding the central idea of visualizing the user's trail on the web. They are different in visualization type, level of interactivity, and the level of user's involvement in the visualization. For example, CZWeb visualizes the visited documents with similar domain names in clusters and allows the user to move the documents between clusters to better match with his mental model (G. Collaud, J. Dill, C. V. Jones, & P. Tan, 1995; B. Fisher et al., 1997). PadPrints is another similar tool that provides a visualization of history map for helping in navigating the web and finding the previously visited web pages in less time

(Hightower et al., 1998). FootPrints are a series of tools including a map of user navigation of the web that has employed the interaction history for contextualizing the web pages to address the problem of navigation in complex information spaces (Wexelblat & Maes, 1999).

In the field of adaptive hypermedia, there is a specific group of technologies called adaptive navigation support systems for supporting user navigation in "virtual spaces" by adapting to the goals, preferences, and knowledge of the user (Brusilovsky, 2004). The studies on adaptive navigation support systems provide insights into supporting navigation of information spaces and thus can inform the design of exploratory search support systems. However, some of their strategies for adaptation cannot be simply used in exploratory search tasks. The examples of those strategies are direct guidance and hiding. Direct guidance by the system may not be possible in exploratory search as the exploration's goal is ill defined and ambiguous for the user. In the hiding strategy, the unrelated links are removed for reducing the cognitive load, but in an exploratory search, evaluating the relevance of links is challenging even for the searcher. However, some of their strategies such as adaptive annotations can be employed in exploratory search support systems. Adaptive annotations augment the links with additional information in the form of visual cues to facilitate the decision-making during navigation, which can be appropriated for exploratory search support systems.

Capturing and visualizing the analysis process has been investigated in some of the visual analytics tools such as CZSaw (Kadivar et al., 2009) mostly

with the purpose of supporting the reuse of the process. Pedersen et al. have proposed the concept of research trails to help the researchers contextualize a fragmented research process without involving the user in organizing the materials (Pedersen et al., 2010). Research trails are ordered sequences of web pages that have been accessed during a fragmented research process; they are automatically constructed based on the user's activity using a combination of semantic and activity based criteria for organizing visited web pages. Both CZSaw and research trails are capturing the exploration process from the system's perspective, that is the captured processes are just the ones that the system has observed and it may not match well with the user's goals and intentions behind each step in the exploration. If the goal of capturing the exploration process is showing the process back to the user for supporting sensemaking and analysis process, it can be considerably enhanced by reflecting the user's intentions behind each action as well as the meaning of each step in the process to the user.

PageLinker (Tabard et al., 2007) is a Firefox extension that contextualizes navigation by linking the web pages when user moves back and forth between web pages by copying and pasting between them. PageLinker recognizes the relevance between pages based on the user's actions rather than using context or domain similarities between pages. Nestor (Eklund, Sawers, & Zeiliger, 1999) is a web browser that draws an interactive webmap while the user is surfing the web. There are similarities between Explore-It and Nestor in terms of the goals and the overall approach for achieving them. Nestor is aimed at facilitating Web

navigation through promoting constructivist learning and supporting reflection in action (Zeiliger, 1998; Zeiliger, Belisle, & Cerratto, 1999). WebView provides a thumbnail-based web history that can be ordered chronologically or hierarchically by site (Kellar, Watters, & Inkpen, 2007). The Hunter Gatherer system allows users to highlight components of a web page and place those components into a persistent notebook using a shortcut key (schraefel, Zhu, Modjeska, Wigdor, & Zhao, 2002).

Explore-It is built upon the previous works and it differs in its focus on providing support for the *process* of sensemaking of the information space under exploration versus facilitating navigation in a domain. We proposed a multi-threaded model for exploratory search process and used this model as the basis of the design of user's trail visualization. The visualization represents the search processes in the user's mind and allows the user to organize his mind and his search process to better make sense of the process and the information space.

4: DESIGN CONSIDERATIONS

In this chapter, guidelines and lessons learned from the theories underlying this work including distributed cognition theory and Schön's theory of reflection in action are discussed.

4.1 Guidelines and Lessons Learned from Distributed Cognition Theory

Distributed cognition is a cognitive theory, which seeks to understand the organization of cognitive systems. The difference between distributed cognition and the traditional cognitive theories lies in their unit of analysis. Distributed cognition goes beyond the individual and considers the artefacts in the environment and the interactions as elements of the cognitive system for understanding the cognitive processes (Hutchins, 1995). It provides a framework for understanding how people interact with their environments and how the environment can assist the user to accomplish highly complex tasks, which are usually beyond the ability of unassisted individual. For this to occur, we need to investigate what external resources can be made available to the individuals and how the distribution of cognitive processes can be established between the internal resources of the individual and the external resources of the environment.

4.1.1 Distribution of cognitive processes in Explore-It: Remembering, and Reflection

There are three possible kinds of distribution of cognitive processes in a distributed cognitive system (Hutchins, 1995). The distribution can be among members of a social group, it may involve internal and external structures, or it can be distributed through time. The distribution of the cognitive processes such as remembering and reflection that are involved in the sensemaking process can be explained by distributed cognition theory. The representation of resources encodes the history of interaction with those resources. An informative representation of a resource, together with the representations of that resource's neighbours building the context around that resource in the visualization, makes it easy for the user to remember a particular resource. Therefore, looking through the lens of distributed cognition theory, the remembering process is distributed among the individual, representations, and the resources per se. This can inform the design of the representations. For example, the representations can be designed to embed the valuable information of a resource that worth remembering. Examples of the information that worth being embedded in the resource's representation are the number of visits to a resource, the time spent on that resource, and how it has been organized in threads or to what specific task it has been related.

The other cognitive process, whose distribution among the elements of Explore-It is discussed, is reflection. Reflection is defined as a cognitive activity for monitoring, evaluating, and modifying one's thinking process. This cognitive activity is distributed among the individual who reflects, the object of reflection on

which the individual reflects, and the surprising factors or unintended results that trigger this cognitive process. By knowing the elements, among which the cognitive process of reflection is distributed, we can design a better system supporting reflection in exploratory search that can ultimately result in a more effective exploratory search. For example, manipulating the objects of reflection so that they call for reflection can be a useful guideline in the designing for reflection.

4.1.2 Externalizing, organizing, spatially arranging the thinking processes

Reflection involves thinking about one's own thinking by monitoring, evaluating, and modifying the thinking process. According to one of the core principles of distributed cognition theory (Boland Jr, Tenkasi, & Te'eni, 1994a; Hollan et al., 2000; Hutchins, 1995), people off-load the cognitive effort to the environment whenever practical. Off-loading the cognitive processes such as thinking can significantly improve the performance in a cognitively challenging task. Reflection can also be enhanced through off-loading the thinking process. Explore-It employs this principle by allowing the user to externalize her thinking process while navigating the Web. This externalization is basically applying the internal organization of the explored web pages that exists in user's mind to the external visualization of the resources by assigning them to different threads, which are representations of the user's though streams. Once the thought streams are externalized, monitoring and evaluating the thinking process for modifying them becomes easier to perform as they have visual representations. Modifying the thinking process may involve organizing thoughts in one's mind

that can be explained by its analogy to spatial arrangement of everyday objects. People are constantly organizing the artefacts in their workspace to enhance their performance (Hollan et al., 2000). The way they manage the spatial arrangement of the items reflects the way they think, behave, and plan (D. Kirsh, 1995). Kirsh proposed three main categories of spatial arrangement based on the purpose of the arrangement: spatial arrangement that simplify choice, perception, and internal computation. The function of arrangement can encode a key piece of information about the problem space. Kirsh's study on how people think about every day activity revealed that people re-arrange items to make it easy to track the state of the task, notice the properties signalling what to do next, and predict the effects of their actions.

For example, Explore-It supports spatial arrangement of resources by allowing organizing web pages in threads that simplifies the perception of relatedness of a resource to an activity with a specific goal. The web pages in a thread are forming a sub-tree in the whole graph of user's navigation trail in which all the nodes are assigned a particular colour that the user has chosen for that thread. Therefore, a sub-tree with the nodes of all the same colour has encoded a key piece of information about the problem space: all the resources in this thread are related to a specific activity with a particular goal, which is the thread name. Visualized record of user's navigation is supporting tracking of the state of the task in general. Also, the colour of a node's border is a property signalling what to do next. The web pages that have been opened but have not yet been visited have orange borders to make it easier for the user to resume her

activity. The orange border is signalling that the corresponding web page has not yet been visited and it can be a potential next web page to visit.

The distributed cognition's perspective on re-arranging the items of the environment with the purpose of simplifying perception, choice, or internal computation can also be explained with the concept of affordance. When we re-arrange the environment, we are manipulating the environment's capacity to make it better lend itself to be used in a desired way. In fact, the user is shaping the affordances of the environment to make the environment afford a specific task, which is required for reaching to a goal. For example, re-arranging some items in the environment with the purpose of simplifying the task of choosing between those items is shaping the environment's affordance for choosing. It can also be considered similar to the design situation suggested by Schön, in which both reshaping the situation and the environment's back-talk are actually revealing the new affordances of the environment, which affords the next moves of the designer.

4.1.3 Appropriating the environment for exploration

A well-designed environment is well pre-structured with rich informational structures such as informative cues to make it easy to deal with the cognitive demands of the tasks in that environment. The activities such as exploration, which is inherently ill structured, demand for highly well structured environment, which can compensate for the complexity of the task. The complexity of the exploratory search is due to several factors, one of which is the lack of a concrete or semi-concrete plan for exploration that is due to the unclear goal

state of the task. Therefore, in order to support effective exploration leading to better sensemaking of a knowledge domain, the environment needs to be designed in such a way that can afford planning, which can contributes to the sensemaking process.

The focus of planning has been on the temporal ordering for organizing the tasks and activities, while the spatial arrangement has been neglected. For instance, the stack model of the structure of goal memory in the information foraging cognitive model is just considering the temporal ordering of the user's intentions. According to ACT-R cognitive architecture, the last item pushed onto the stack is the first item to look at and this shows the temporal ordering. However, Kirsh has revealed the everyday cases in which people are using spatial arrangement to plan for their actions (D. Kirsh, 1995). Considering spatial arrangement as an important factor in planning, the environment needs to afford arrangement. Exploration Path, one of the Explore-It's components, affords spatial arrangement of its elements. Nodes, representations of web pages, in Explore-It can be re-arranged to support planning. Exploration Path allows the user to move the nodes between the threads (Figure 2).
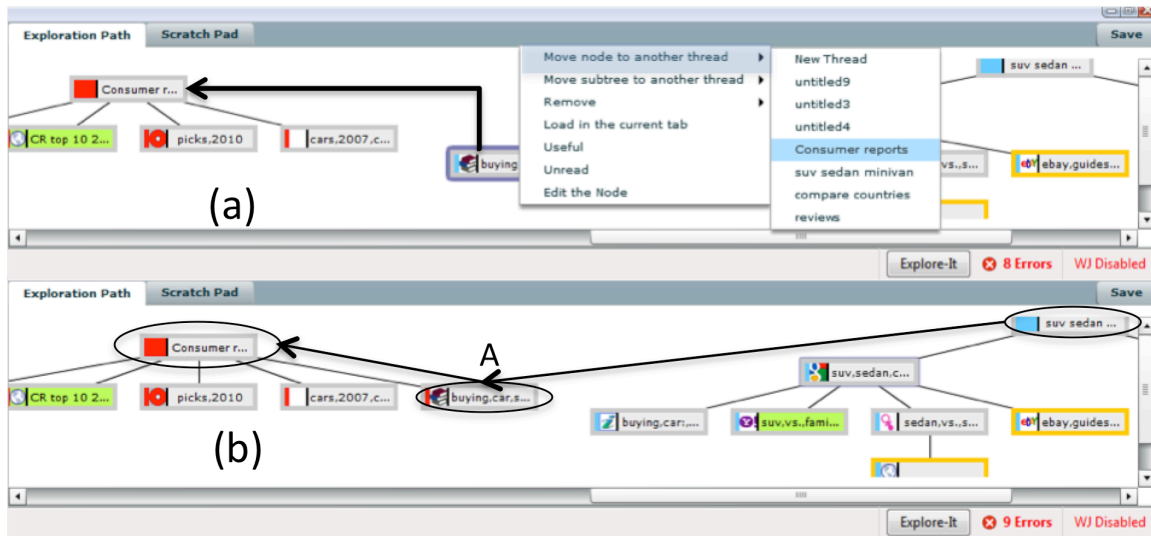
Figure 2: Re-arranging the Exploration Threads' structures by moving a node from one thread to another. (a) Illustrates the scenario of selecting a node from 'suv sedan' thread and choosing to move that to 'consumer report' thread; top figure illustrates the scenario of selecting a node to move that to another thread, whose name can be found in the context menu. (b) Re-arranged threads; node 'A' has moved from one thread to another.

4.1.4 Design of representations in Explore-It

The nature of representations and the way people work with them is a key focus of distributed cognition. Hollan et al.'s cognitive ethnographies have shown that people often shift their attentions back and forth between the properties of the representation and the properties of the things being represented. This strategy of shifting in and out of the representation has been observed to support some cognitive processes. We are interested to know how users interact with the Firefox tabs and their representations in Explore-It; and how they are shifting their attention between those two. The nodes in the graph visualization of the user's trail in Explore-It represent the Firefox tabs (Figure 3).
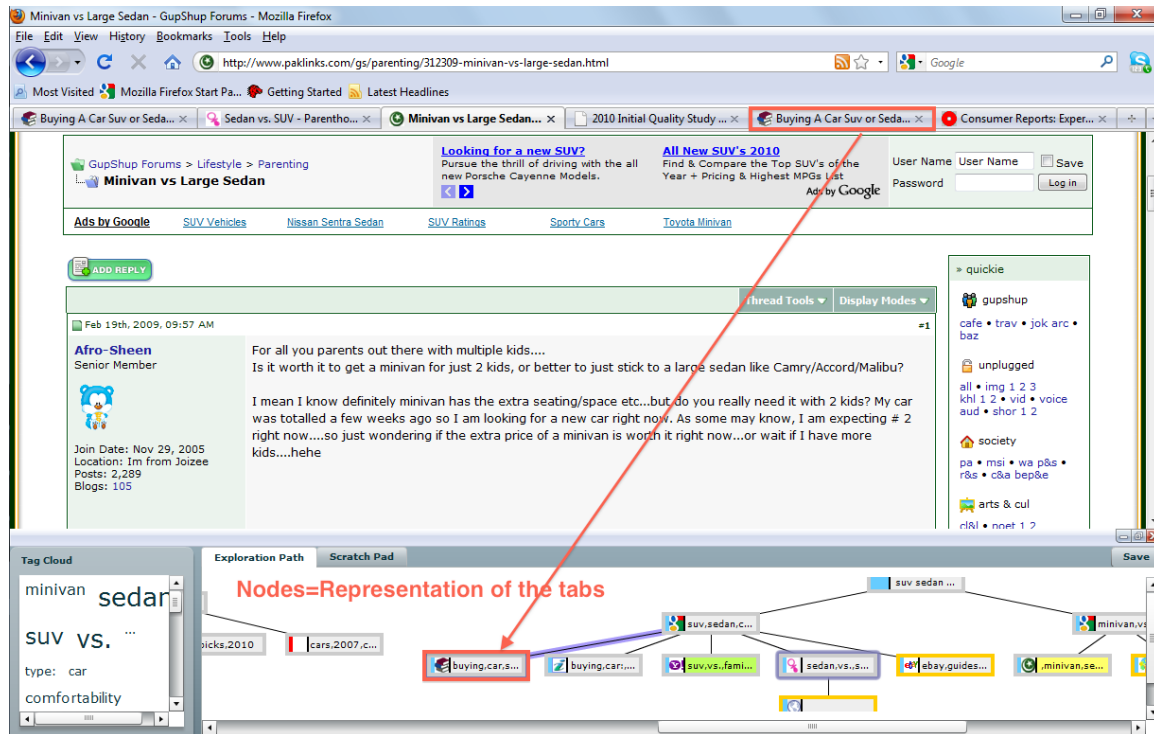
Figure 3: Nodes in the graph visualization represent the Web browser's tabs

The node in the Exploration Path has been designed to represent a summary of visual and textual information of a web page. The design of a web page's representation is fully described in (5.2.3).

Construction and communication of sensemaking representations: A critical aspect of sensemaking is finding appropriate representations to the information important to the task at hand (Qu & Furnas, 2005) that can enrich the communication of the sensemaking outcomes to the self as well as to others. A system aimed at supporting distributed cognition should allow individual to build representation of the context in the process of constructing interpretations and to exchange the representation with others (Boland Jr, Tenkasi, & Te'eni, 1994b).

Embedding the context in the representation of sensemaking outcome enriches communication of the sensemaking outcome with the self and others. People come up with specific sensemaking representations appropriate for a particular task type (Qu & Furnas, 2005). In order to make the representations communicable, the system should capture the possible reasons behind construction of those representations by the user. For example, one of the sensemaking representations in Explore-It is the Exploration Thread, which is constructed in a mixed initiative approach by both the user and the system. Both the user and the system encode the context of user's activity in the Exploration Threads; user by choosing a related name for a thread, assigning tags to a thread, and assigning web pages to the threads and the system by tracking the user's path from one web page to another.

In Explore-It, the tag could of user's interest is capable of providing more context to the threads. Therefore, it makes the threads more understandable, when user is resuming the search after a long interruption. The tag cloud shows keywords extracted from several search related activities such as queries, tags, threads' names, comments on nodes, and web page's titles.

4.2 Guidelines from Schön's Reflection in/on Action

Reflection is defined as a cognitive activity for monitoring, evaluating, and modifying one's thinking process. The notion of reflection-in-action, introduced by Donald Schön (Schön, 1983), is sometimes described as 'thinking on our feet' and it involves looking back to one's experience and attending to one's theories

in use. It entails building new understandings to inform our actions in the situation that is unfolding and simply means thinking about the action while doing it.

One of the goals of Explore-It is mediating reflection-in-action in the exploratory web search, i.e. encouraging user's thinking about what they are doing and evolving the way they are doing a specific action during the process. However, reflection-in-action is not simply happening in every circumstance rather is associated with the experience of surprise. According to Schön, when intuitive and spontaneous action leads to expected results, we usually tend not to think about that action. Instead, unwanted and surprising results of an action might be responded by reflection-in-action (Schön, 1983). However, the unexpected results can induce reflection-in-action just for the ones that are somehow engaged in the action, while they care about the results of their action and are motivated to improve the way they are doing that action. The potential for reflection-in-action can be predicted by the nature of the result of the action, whether it is expected or surprising. There are some actions in which the users do not have any expectation of what the result should be. The example of this kind of action is exploratory search, in which the user lacks knowledge of the domain. The user can agree with every argument of every source that she encounters during her exploration, as there is no expectation of the result of the search. In order to avoid designing the system just for a specific type of users, who are interested in improving their search processes and are eager to learn more about the domain under exploration, we considered *inducing* reflection-in-action in addition to *supporting* reflection-in-action. Before describing how

Explore-It induces or supports reflection in/on action, we discuss the possible objects of reflection, the processes or products on/in which the user reflects.

4.2.1 Object of reflection

The practitioner can reflect in/on several things. She can reflect on her implicit strategies in her activity, on a feeling for a situation that led her to do a particular action, or on the way that she had framed the problem for solving it. The objects of reflection can be divided into two general categories of products and processes. According to Lin et al. (Lin, Hmelo, Kinzer, & Secules, 1999), reflection should occur with at least two objects of reflection, the product and its value, and the process. Reflecting on the process is generally hard to support, as the process is hidden and sometimes difficult to track. In contrast, it is much easier to support reflection on the products, as they are the explicit results of a process. Explore-It provides support for both categories of objects of reflection, the product and the process. The products on which the user can reflect in an exploratory search activity are the resources that she finds and her notes or comments on the web pages. The processes include the overall process of exploration and all the user's action steps such as opening a web page, closing a web page, and etc.

4.2.2 Mediating Reflection in/on Action in Explore-It

Following is a description of how Explore-It is designed for inducing and mediating reflection in/on action.

1. Assigning a web page to a thread: The way the user is involved in the organization of her Exploration Path is designed to provide the user with the opportunity to reflect on the process. The user is given the option of naming a search thread and assigning a web page to a thread. Although the system is implicitly asking the user to think about what she is doing and how the new web page can be related to her task, it is not prompting reflection. The user is still allowed not to think about her action and proceed in her search by ignoring the system message.

2. Visualizing of the process: The node-link graph of the user's trails in her search is not supposed to just represent the history of browsing. It also represents the user's thought streams via the Exploration Threads. Visualization of the process can induce reflection in action as well as reflection on action. It shows an overview of the process as well as the connections between different steps in the process that could hardly been attended when the user is engaged in each single step.

3. Providing writing tools: Once the user is induced to reflect, there should be enough support for her to be able to externalize her thoughts. As she externalizes her modified understanding of the situation, she can better internalize the patterns emerging from the aggregation of previously existing information and the recently added information. Therefore, Explore-It provides annotation tools to allow the user to modify the properties of web pages' representations such as the thread they belong, and to write note for web pages.

4.2.3 Effect of reflection on learning

Bannert investigated the effect of reflection on learning with hypermedia using metacognitive prompts (Bannert, 2006). Metacognitive prompts are basically questions that make the user think about her action. Bannert designed a metacognitive prompt to make the students reflect on their navigation behaviour by asking them to say aloud the reason of selecting a specific node. Inducing reflection by metacognitive prompt has been proved to be effective on resulting in better outcome in several studies. However, we agree with a view of reflection as a fully engaged interaction as opposed to a detached assessment of the process or its product. Explore-It has integrated support for reflection-in-action into the user's actions in the process of searching. Inducing reflection in an implicit way does this integration.

Lin et al. identifies four types of design features that scaffold for reflective thinking in a learning environment. These scaffolds include process displays, process prompts, process models, and reflective social discourse (Lin et al., 1999). "Process displays" are used for displaying problem solving and thinking processes. "Process prompts" are designed to prompt student's attention to specific parts of the process while learning. "Process models" are aimed at providing models of experts' thinking processes, which is usually tacit and cannot be communicated by the experts themselves. The goal for modelling experts' thinking processes is to allow the students compare those processes with their own processes and reflect on them. Lastly, "reflective social discourse" is about creating community-based discourse to access variety of perspectives and

feedback, which are reflection-inducing factors. The core idea of all these design features is visualizing the thinking process, which is one of our main goals that we hope to achieve it by designing Exploration Thread capable of visualizing the user's thought process.

4.2.4 Reflection and memory

There are some evidences revealing the shortcomings of reflection that arise from the nature of memory (R. Newell, 1992). In order to reflect, one should reconstruct her experience and this requires the cognitive process of recalling. Reflection in action involves immediate recall and reflection on action involves delayed recall. While both heavily rely on the memory, they differ in the time between acquisition and retrieval of a memory. Some people believe that reflection is a flawed process because of its reliance on memory. However, this fact informs the designers to consider support for memory and recall when they are designing for reflection. There is ample evidence that recall is linked to variables associated with the acquisition of memory (J. R. Anderson & Bower, 1980). Therefore, recall can be enhanced, if it happens in a situation similar to the one in which acquisition took place. Taking these facts to the design context, having enough action's history embedded in the situation can enhance the reflection by triggering the memory through showing the enriched context of action.

4.2.5 Sensemaking: a reflective conversation with the information space

Sensemaking can be described by analogy to design; it can basically be considered as a design process in which the sensemaker is involved in shaping her understanding as she encounters new information and situation and creates intermediate artefacts as a result of that understanding. These intermediate artefacts are the sensemaking representations such as concept maps, notes, and etc. Therefore, a sensemaker, who is making sense of a knowledge domain through exploratory web search, is performing design-like tasks, as she is constructing representations. The sensemaking representations are the outcome of the sensemaking process representing the user's understanding of the domain.

Regarding Schön's "back-talk" metaphor, when the designer applies some changes to the situation, the situation talks back to the designer and the designer responds to the situation's back talk. Likewise, using the analogy between design and sensemaking, sensemaker in the exploratory search can apply her understanding of the domain under exploration to the visualization of the path that she has taken. The changes that a sensemaker is making to the situation are in the form of organizing, moving resources in different categories, and annotating the nodes of the visualization. In fact, depending on what we are considering as a situation in the case of sensemaking, there can be various changes to the situation. We define 'situation', in the context of exploratory web search, as part of the information space with which the user is dealing, including the resources, web browser and different parts of Explore-It as a web browser

extension. Therefore, changes to the situation can be considered as any changes to the web browser contents, the Exploration Path and the tag could. We clarify this analogy by describing the examples of exploratory search and interpreting the situation, the change to the situation, the situation's back-talk, and the user's response to the situation back-talk for those examples.

Example 1: Modifying a search query. Considering a web search query as a part of the situation of sensemaking, any changes that the user is making to the query is changing the situation too. The user is shaping her view of the information space by tailoring the whole information space through her search query. The search engine's results are the situation's back-talk to the user's changes to the query. The user responds to the situation's back-talk by selecting specific pages to visit or refining the query with the new understanding gained from the situation.

Example 2: Changes to the visualization. When the user makes any changes in the graph visualization of the user's trail such as moving nodes from one thread to another, she is changing the situation in which she is making sense of a particular domain. The modified graph talks back to the user as part of the situation that might reveal some other connections between the nodes. The user further can take some actions based on the new information that the visualization has revealed to her. These actions can be creating a new query, leaving note on some nodes, finding out about other existing relations between resources, and etc (Figure 2).

The analogy between sensemaking and design is apparent in all the above examples. The following scenario was common in both cases:

- The user is reshaping the information space to satisfy her information need

- The reshaped information space talks back to the user while providing her with more information

- The user uses that back-talk in her further actions

This is exactly similar to the design and what Schön proposed as the reflective conversation with the situation (Schön, 1983).

4.2.6 Challenging Schön's reflection in action

The way Schön describes the situations in which reflection can occur seems to be restricting the areas in which reflection can be beneficial. However, we should distinguish between when or where it's likely for the reflection to happen and when inducing reflection can enrich the way people act in a specific situation and hence is beneficial. According to Schön, 'surprise' is a strong factor in reflection and it usually occurs when there is some puzzling, troubling, or interesting phenomenon with which the individual is trying to deal (Schön, 1983). There are other circumstances in which no element of surprise can be found, however, reflection can be helpful as there is always room for improvement even if the result of reflection does not directly influence the current action. Therefore, we should seek ways to support both types of reflections, with or without the surprise element.

Although we believe that *reflection-on-action* can always be beneficial, *reflection-in-action* cannot be guaranteed to be useful for the current action as it strays the user from the *flow* of his action. Flow is the mental state of operation in which the person is fully immersed in what she is doing (Csikszentmihalyi, 1998). There seems to be a trade-off between staying in the flow and reflecting-in-action. The skill level of the person and the challenge level of the activity characterize the flow state; if a person is engaged in a challenging activity with a set of high level skills relevant to the activity, it is likely for her to fully immerse in that activity. These prerequisites for being in the flow are the same prerequisites for reflection to occur. However, these two seem to be contradictory to each other, as reflection requires monitoring and evaluating what you are doing which is not happening in the flow state. Therefore, in designing for reflection, the designers should consider the costs associated with reflection that is "not experiencing the flow of activity".

5: EXPLORE-IT: AN EXPLORATORY WEB SEARCH SUPPORTING TOOL

The primary goal of Explore-It is to support exploratory web search by facilitating sensemaking of the information and sensemaking of the exploration process. In order to achieve this goal, we incorporated the design guidelines from (chapter 4:Design considerations) into designing Explore-It's component including the tag cloud and Exploration Path with all its elements. Before discussing the components of the system, we will discuss how the primary goal of our design is broken into several sub-goals, which are then used to form the rationales behind the design of components of the system.

The target task corresponding to the sensemaking activity in an exploratory search is not clearly defined, especially early in the process. At first, the target task can be considered as finding the goal, which can further determine the next target tasks. Regardless of having a specific target task or not, the sensemaking activity has still its own sub-activities or actions that should be supported. According to Russell's sensemaking model (Russell et al., 1993), the common actions in every sensemaking activities are creating representations, encoding information into these representations, and using the structured information at the service of the target task.

Enabling the user to organize her view of the domain under exploration is essential for supporting construction of representations (Brian Fisher et al.,

1997). However, asking the user to organize her web resources clearly increases

the user's workload; especially, when the user is involved in exploratory search,

in which he cannot create appropriate representations due to the lack of domain

knowledge. On the other hand, when the user has no idea of what an appropriate

structure would be for organizing the information, the system will have a hard

time in creating that structure. Therefore, we decided to use a mixed initiative

approach to creating a simple representation of the web pages that has the

following characteristics:

1. Close to the user's mental model of the domain,

2. Reflects the user's sub-goals and thought streams that have been
   followed,

3. Easily modifiable so that modifying of the representation does not take
   too much time and effort.

Finally, in order to make the user aware of the changes in her interest

throughout the exploratory search process, Explore-It provides a tag cloud

containing the keywords extracted from different activities of the user. Each of

these components is discussed in the following sections in more details.

5.1 The Concept of Exploration Thread

We developed the concept of "Exploration Thread" as the representation of

the user's thought stream around one aspect of the information need. According

to sensemaking models' terminology, we have considered each thread as a

representation that can organize the information related to the exploratory search

process. Organized information can reduce the cost of making sense of a target information task, which can be gaining a general understanding of a knowledge domain. For seeing some examples of the Exploration Threads, please see Appendix 4.

We seek to understand to what extent these threads can represent the internal sensemaking representations so that we can consider them as externalized sensemaking representations—representations that are constructed completely by the users to serve an output task such as writing a report or preparing a talk. In addition to serving an output task, Exploration Thread has been designed to serve the process of exploratory search per se. There has not been much elaboration on the type of the organization that the user needs to organize the exploration process for having more control over the process. Sensemaking of the exploration process mediates the information foraging phase and the sensemaking of data.

5.2 Exploration Path: A Visualization of the User's Trails

Exploration Path is a graph visualization of the user's trail while exploring the web (Figure 6). An external representation of the search process can help the user in several ways: 1. Guiding the exploratory search by providing awareness of the whole process, 2. Allowing off-load of multiple concurrent ongoing search threads in user's mind to avoid switching back and forth between different threads with the purpose of not losing track of any of them, 3. Enabling communication of the information need in a broader context through sharing the process, encoded in the Exploration Path, 4. Enabling more contextualized and

easier-to-understand recommendations on the exploration by integrating them into the visualization.

Keeping the history of explorations is a critical consideration in designing exploratory search support systems, as usually searchers need to go back and forth between their findings to get a sense of the domain under explorations. The visualization is a simple node-link graph in which the nodes represent web pages, visited or planned to be visited. If visiting node 'A' introduces node 'B' to the searcher and the searcher continues her exploration from 'A' to 'B', there will be a link from 'A' to 'B' in the Exploration Path. Visiting a page, whether it's a new article or a web page, can introduce new references, keywords, authors, journals, and conferences, all of which are shaping the next steps of the exploration and are the causes of visiting the next series of pages. The graph is created while user is exploring a domain by searching and browsing web pages. It is constructed with a mixed initiative approach, in which the user can be involved in organizing the graph's nodes according to her goals and sub-goals. *Exploration Threads* are representing the user's path while reflecting user's sub-goals in the exploration. Exploration Threads are meant to demonstrate the pathways that the user is traversing to achieve a certain sub-goal. However, the sub-goals can hardly be recognized by the system unless by looking at user's actions from the user-situation viewpoint. This way, Exploration Threads can also represent user's thoughts about what's happening and what does it mean to the user (Albers, 2005). Mixed initiative approach to visualization has been one of the suggestions for applying human cognition model for visual analytics (Green, Ribarsky, & Brian

Fisher, 2009). Employing a mixed initiative approach to construction of the visualization, Explore-It is providing a basic graph visualization of user's trail and user can restructure the graph based on her own understanding. In addition, the user can enrich the visual representations of the web pages and her thought streams by assigning tags to the nodes and threads.

5.2.1 Understanding and Modelling Exploratory Search Behaviour

In order to understand the search behaviour of users in an exploratory search, we decided to focus on researchers in the early stages of research who are searching for related literature on a topic to gain a comprehensive understanding of a field. We did a pilot study performing an exploratory search while keeping track of our thoughts by writing notes to get a better sense of the user's need in exploratory search. In addition to that, we had informal formative interviews with two graduate students and we asked them to explain their search behaviours as well as their needs in exploring a new knowledge domain outside the area of their expertise. We found that exploratory search behaviour can in part be explained by SNIF-ACT (Fu & Peter Pirolli, 2007), which is a model of user's navigation on the web based on the ACT-R cognitive architecture described in (2.3: Cognitive Architectures). However, the major difference between the information seeking behaviour explained by SNIF-ACT model and the exploratory search behaviour that we found, based on the pilot study and the informal interviews, was the structure of the user's goal memory. The goal memory, hierarchy of user's intentions driving the information seeking behaviours, is modelled as a push stack in SNIF-ACT. However, we observed

exploratory search process as a multi-threaded process in which the user is

following multiple sub-goals, all of which are contributing to making sense of the

domain (Figure 4). The stack model of goal memory cannot explain every type of

user's goals in their information seeking behaviour. There are some sub-goals

that can be followed in parallel, although not necessarily concurrently. The

parallel nature of sub-goals implies the fact that they are not necessarily

dependent on each other in the sense that achieving a sub-goal requires

achievement of another sub-goal. The user follows a seemingly independent

sub-goal in each search thread.

(a) goal memory structure in
exploratory search

(b) Stack goal memory
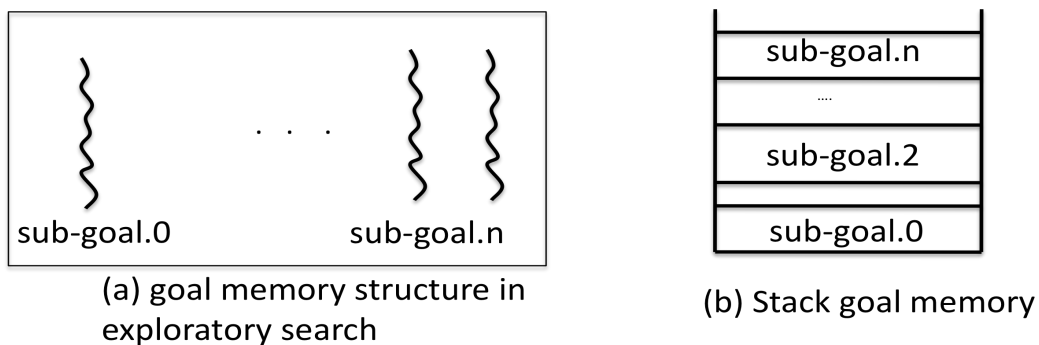
**Figure 4: The structure of user's goal memory a) in exploratory search, b) in SNIF-ACT
cognitive model**

5.2.2 Multi-Threaded Exploratory Search Model: An Example

This model of exploratory search can be clarified by the following

example:

John wants to buy a digital camera and he is not familiar with photography

and digital cameras to the extent to which he doesn't know what are the salient

features of a camera to look for. John begins his search with submitting "best

camera" as a query to a search engine and continues with visiting some of the resources he found in the results' page. While visiting those web pages, he realizes that there are many terms that he does not know their meanings. Examples of those terms are DSLR, compact, live view, and etc. First, he decides to understand what is DSLR or compact and what is the difference between the cameras characterized by these two terms. In this step, understanding these two terms and the differences between them becomes a sub-goal in exploratory search. While reviewing the results of the new search thread, he comes across the "micro four thirds standard" and he starts a new search for knowing more about that standard. As he continues in either of the previous search threads, he remembers his interest in bird photography; therefore, it becomes important for him to get an appropriate camera for bird photography. So, he starts a new search thread by searching for "bird photography camera" in order to know about the cameras' specifications important to bird photography. The sub-goal in this thread is "understanding the important features of a camera for bird photography". From the search results, he learns that the specifications of cameras' lenses are important in bird photography. However, he has difficulty in understanding different camera lenses' types, as he is not familiar with the characteristics of camera lens including "focal length", aperture, "depth of field", and etc. So, he begins a new search thread with the sub-goal of "understanding camera lenses" which will eventually results in creating a number of sub-threads in that thread with their

own sub-goals each associated with understanding different characteristics of camera lenses.

In this example, the threads are not dependant on each other, meaning that completion of one requires completion of a number of them. The user can move between threads and keep multiple threads active at a time. In an exploratory web search, the task requirements are not clear at the early stages of the task. As new requirements for the task are uncovered through exploration, new sub-goals are added as new Exploration Threads to the Exploration Path. However, the process can be completely different for a professional photographer who has a comprehensive knowledge of cameras. Bhavnani observed experts performing search tasks within and outside their domains of expertise. He found that knowledge of goal sequencing and knowledge of specific URLs for each goal is specific to experts who were performing search within their domains of expertise and that influenced the efficiency and effectiveness of their search (Bhavnani, 2002). Non-expert information seekers can rarely have a particular goal sequence in mind while exploring a new domain outside their expertise (Kang & Fu, 2010; Bhavnani, 2002). Explore-It tries to support goal sequencing for non-expert searchers by facilitating reflection in exploration in order to develop knowledge of goal sequencing through reflection.

5.2.3 Design

In this section, we will describe the rationales behind the design decisions in designing Exploration Path (Figure 6).

Visual representation of web pages

The nodes of the graph visualization in Exploration Path are representing web pages; all the elements of Explore-It including the nodes in Exploration Path have been designed with the goal of supporting sensemaking through facilitating self-reflection and communication of information needs to others if necessary. Visiting a page might inspire the user to reflect on another page that has been visited before. In order to make it easier for the user to find the node representing a web page to comment on that or to assign some tags to that, the node should remind the user of that page without overloading him with unnecessary information of those pages. We call this design principle "Just Noticeable Reminder". Having this principle in mind, the following studies have also informed the design of the nodes in the Exploration Path. A study on what people recall about their documents in a PC has shown the important role of the visual elements in recalling a document (Blanc-Brude & Scapin, 2007). Consequently, several studies have been informed by this finding and have used thumbnail of visited pages to help the user in revisiting a page from the graphic history of their browsing (Ayers & Stasko, 1995). However, visual snippets and enhanced thumbnail (Woodruff, Faulring, Rosenholtz, Morrsion, & Peter Pirolli, 2001) that consists of a salient image and some keywords of the webpage can enhance the recognition more than thumbnails and textual snippets (Teevan et al., 2009). Also, considering the limited screen real estate, thumbnails can hardly be useful in relatively small sizes. Being inspired by all the aforementioned findings in the

previous studies, we considered encoding both visual and textual information in

the nodes that are representing web pages. The nodes in Exploration Path are

composed of three parts: 1. A narrow coloured bar indicating the theme of the

thread that has been determined by the user for distinguishing different threads,

2. Web page's icon, the tiny icon displayed on the browser tab for that web page

(called Favicon), 3. Comma separated keywords extracted from the html title of

the web page. Instead of extracting the most salient image of the page, we

decided to use the web page's icon to minimize the error in extracting the most

salient image as well as saving the screen real estate. The coloured bar is

included to segregate the Exploration Threads and give an overview thereof

without having to look at the tags that have been assigned to the thread. A

sample node is shown in (Figure 5).
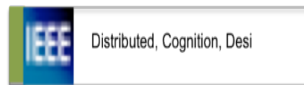
Figure 5: Sample node in Exploration Path consisting of a colored bar, the web page favicon, and the keywords extracted from the title of the web page

Flagging of useful pages is integrated directly into the representation of the pages in the visualization.
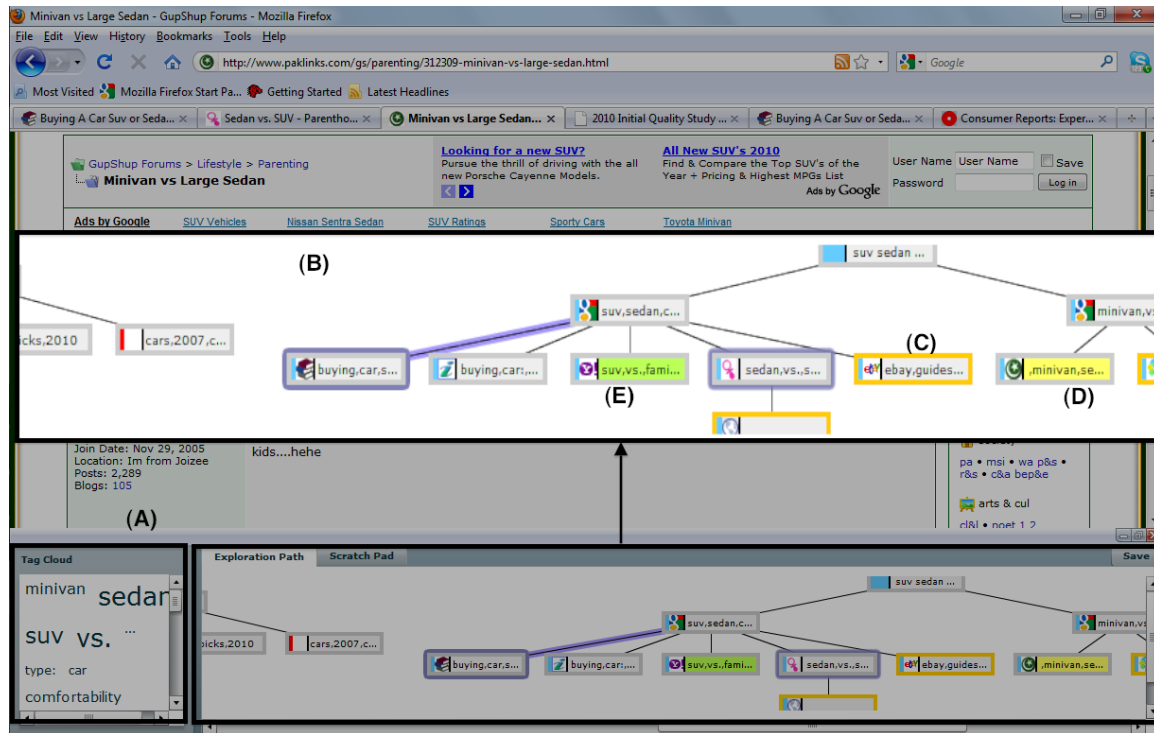
Figure 6: Explore-It system: (A) Tagcloud: illustrating the keywords extracted from the user's query, web page's urls, and thread names and Web page's comments (B) Exploration Path: a graph visualization of the user's trail in exploring the World Wide Web (C) Node with an orange border are not yet visited by the user (D) Node with the yellow background shows the current tab at which the user is looking (E) Node with the green background shows that the user has marked the node as 'useful'

Visited web pages are distinguished from the opened and not yet visited ones in the Exploration Path

An ethnographic study of researchers has revealed their need to know which tasks have been remained unfinished or where they have left off last time that they were working on a specific search task (Pedersen et al., 2010). The

nodes representing the web pages that have been opened but have not yet been visited have orange borders to be distinguished from the visited ones (Figure 6)(C). Having a number of open web pages while some of them have not yet been visited usually happens when the user finds many interesting links in a web page. As the user cannot go over all those links simultaneously, she starts opening each of those in a new tab, although she might focus on one of them at the time. Differentiating between the web pages that have been visited and others is especially important when the user is working on the search task in multiple sessions. It can facilitate resuming the task from where it was left off. As the user has spent a certain amount of time on applying a loose structure on the representations of opened web pages, he can simply resume his task from the point he left off without having to re-spend that time remembering which of those pages have been already visited. This also justifies our decision of the time of asking user to assign the selected link to a thread. When the user is clicking on a link, he will be asked to specify whether the new web page belongs to the current thread, a previously explored thread, or a new thread. The user's intention from clicking on a specific link is clearer to him at the time he is deciding to click. However, this intention can be changed once the user visits the page and knows more about the content of the page. This clarity in his intention at that time helps the user in determining a thread to which that page should be assigned.

Showing the corresponding web page's URL

According to a user study (Won, Jin, & J. I. Hong, 2009), most of the users do not remember any URL unless it's short and descriptive of the web page.

Therefore, we decided to show just the descriptive part of the URL, in the tooltip when the mouse hovers a node. The extracted keywords from the URL and from the web page's html title construct the descriptive part of the URL for this purpose. Considering our design principle of "Just Noticeable Reminder", when the user moves the mouse over the node, the descriptive part of the URL will be shown in a tooltip not to overload the user with unnecessary information. However, the nodes are linked to the pages they represent and the user can visit the page without having to know the URL.

Facilitating user's reflection

Reflection enhances learning through enabling understanding gained from one's experiences. Explore-It has been designed to facilitate and encourage reflection through integrating some activities related to reflecting into different steps of the search process. Those activities and the steps of the search process in which they are integrated are summarized in (Table 1).

Steps in the search process	Activities	Facilitating/encouraging reflection
Navigation (opening a hyperlink)	Assigning the web page to a thread before opening it	Encouraging
Gaining a new understanding of either the domain or the structure of the information in that domain	Changing a web page's thread	Facilitating
Finding a resource useful	Marking the node representing that resource as useful	Facilitating
Thinking about what was read and coming up with a new understanding	Writing notes/comments on the node	Facilitating

Table 1: Activities integrated into the different steps of the search process to either support or encourage reflection

While opening a hyperlink, user is required to assign the web page to a thread. This assignment will encourage the user to think about her purpose in opening that hyperlink. This can especially help if the user later revisits that web page, since she may have forgotten the reason she had in mind at the time of opening the hyperlink. Another step in the search is when user has gained a new understanding of either the information or the information structure in the domain,

in which case user might need to apply some changes to the threads' structures such as changing some of the web pages' threads. By allowing the user to apply these changes to the visualization of her trail, Explore-It is facilitating reflection in this step of the search process. Other steps in the search including reaching to a conclusion or finding a resource as useful are accompanied with the activities of writing notes and marking nodes as useful which are facilitating reflection.

The path is providing strong cues in recalling a webpage

Providing context is a strong cue in recalling (Spear, 1978). Visualizing the explored path provides a strong cue in recalling a web page as it's showing 1) the visitations of other related web pages, 2) notes on the web pages, 3) the name of the thread to which a web page has been assigned, and 4) the graphical icon of the web page. All of these are contributing to the distributed cognitive process of reminding.

In sum, the artefacts involved in the cognitive process of recalling a web page and the story behind that are as following: the nodes representing web pages, the edges representing the paths from one web page to the other, and the textual and visual information embedded in the node representing keywords of the web page's title.

Annotating the Nodes of the Exploration Path

The user can reflect on her activities and enrich each graphical node, representing a Web page, with some textual information, which can be the result of sense made from the information or the structure of information. Assigning

tags to the threads for describing the searcher's goal in traversing that thread can enhance information scent (S. Zhang, Farooq, & John M. Carroll, 2009). Information scent refers to detection and use of cues that provide the user with the information that is not immediately available. Information scent cues in the world activate the cognitive structures called chunks (J. R. Anderson & Lebiere, 1998). This activation then spreads from these cognitive structures to the related structures in the activation network. Therefore, the user can have a better understanding of where to go for finding a particular piece of information. More importantly, she will find a better sense of the connections between the fragmented pieces of information.

5.2.4 Organizing the Nodes in Exploration Path: A Mixed Initiative Approach

The exploration path is not completely constructed automatically by the system. When the user clicks on a link, the user will be asked to decide about the thread to which the web page should be assigned. The options include assigning to the current search thread, to one of the previously explored threads, or to a new thread based on the similarities and differences between the sub-goals of those search threads. Therefore, this is the user that determines the location of a new node, representing a web page, in the Exploration Path. However, user's organization of the nodes may soon get outdated as the user's knowledge increases during the exploration. Therefore, Exploration Path is designed to be flexible to these changes by supporting the movement of nodes between threads. Assigning web pages to different threads is the first level of user's involvement in

organizing them. The second level is assigning some keywords/tags to each search thread that can represent the implicit sub-goal of creating a new search thread. Completely imposing the responsibility of organizing web pages on the user has many disadvantages such as increasing the user's cognitive load and distracting her from the primary task of searching. In addition, it has other disadvantages that are particularly associated with the nature of exploratory search including the unclear goal and the expanding knowledge of the user. These characteristics of exploratory search can make the user's organization of web pages inappropriate. Also, it is likely for them to need modifications to be useful for a certain task. Therefore, the mechanism and operations for updating the organization has been designed to allow for fast and easy modifications to the organization.

A bookmark system is an example of systems that imposes all the organizational responsibility on the user. These responsibilities include adding structure to bookmarks by building hierarchies and tagging them. However, most of the bookmark users are not interested in organizing them at all (Abrams, Baecker, & Chignell, 1998). On the other hand, there are systems that never involve user in the organization; examples of such systems are the ones providing visual browsing history of the user (Hightower et al., 1998; Ayers & Stasko, 1995; Doemel, 1995; Frécon & Smith, 1998). These systems are less likely to offer useful structures for the user's task, as they are not open to new structures suggested by the user. Mixed initiative systems are in the middle of this continuum, where both the user and the system are involved in organizing

the explored materials (Figure 7); the system creates an initial graph of the web pages and the user is allowed to change the structure. CZWeb is an example of this type; it creates clusters of visited web pages based on the domain of web pages and then the user is able to move the pages between clusters to better reflect his mental model of the visited web pages (G. Collaud et al., 1995; B. Fisher et al., 1997) (B. Fisher et al., 1997). According to Russel's sensemaking model (Russell et al., 1993; Furnas & Russell, 2005; Qu & Furnas, 2005), mixed initiative systems can be more successful in supporting sensemaking for two reasons. First, they can reduce the cost of structuring the information by providing a basic modifiable structure, not requiring the user to build the structures from scratch. Secondly, they can increase the gain, defined as the increase in quality or quantity of the performed work (Russell et al., 1993), by reducing the time spent on the target tasks such as analysis, preparing reports, and etc., due to the familiar and meaningful organizational schemes to the user as they have been created by himself.
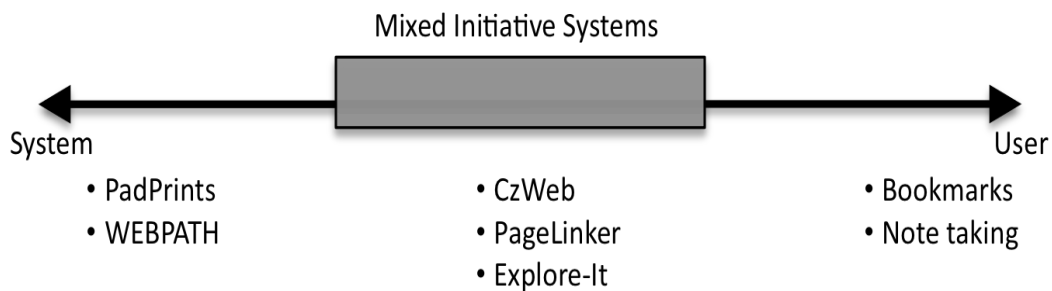
Figure 7: The continuum of user-system involvement in organization of web pages and the examples of the systems in this continuum

The stage of the exploratory search process can determine the level of ambiguity that the user is experiencing in the search. The ambiguity of the domain under exploration should decrease as the user spends more time on exploring the domain; but considering the exploration process in more detail, each new piece of information can increase or decrease the ambiguity of the situation (Figure 8). Therefore, the stage of the exploratory search can be a determining factor in deciding on how much the user and the system can and should be involved in the organization process. In the early stages of the exploratory search, the user is experiencing a high level of ambiguity in the domain. The system's involvement in applying structures on the information space would be useful at this stage as the user does not have to apply loose structures on the information which might also get outdated soon due to the poor knowledge of the domain at early stages. On the other hand, when the user has gained more knowledge of the domain after exploring the domain for a while, she can be more involved in applying appropriate structures, according to her task, on the information space. However, we haven't considered adaptive mixed initiative approach in this thesis and the level of the user's and the system's

involvements do not change during the search in Explore-It.
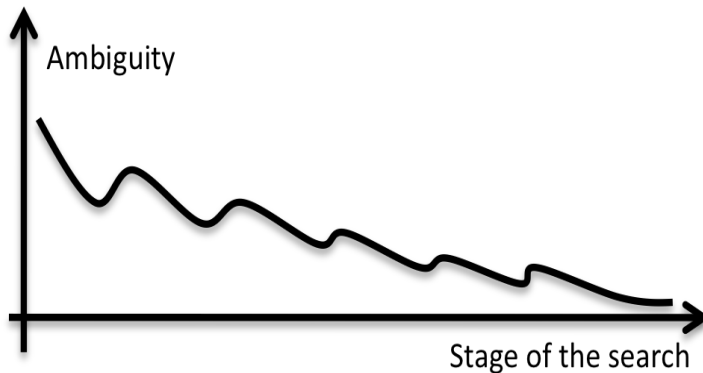
Figure 8: The level of ambiguity that the user experiences in the exploratory search varies with the stage of the user's search

5.3 Tag Cloud: Representing Evolution of User's Interest during the Search

Although Exploration Path is providing awareness of the process for the user by allowing the user to off-load her sub-goals in exploration, it is not an efficient representation of the user's interest. User's interest is constantly evolving during exploratory search. As the user does not have a very explicit goal, she simply follows the new ideas and directions introduced by each new piece of information along his path (Bates, 1989). While being guided through some routes by her selections of the pages, it's likely for the user to forget the origin of the path and the main topic for which she started her search. Therefore, enabling the user to keep track of her interest in the domain is important in an exploratory search support system. Explore-It has a tag cloud, which provides the visual depiction of the words that have been likely to be of interest to the user. We consider a word is of interest to the user if it is used:

59

- In the search query, as the query is the first place that the user is explicitly announcing a shift in his interest,

- As a tag or name that the user has chosen for Exploration Thread,

- In the title of a web page that has been opened,

- In the user's comment on web pages

The words are added to the tag cloud from the top, meaning that the most recent words are in the top of the tag cloud. Additionally, the size of a word in the tag cloud represents how frequently the word has been of interest to the user (Figure 6).

5.4 Usage Scenario

In this section, a usage scenario that illustrates Explore-It in use is described. One of the participants' scenarios has been chosen for illustrating the system's functions. We had an extended interview with this participant and asked him to walk us through some of the steps again for describing the main functions of the system. The participant was the first-time car buyer who had chosen to search for finding the right car for himself. In this scenario, the hypothetical name of this participant is Peter.

There are three ways that a user can visit a web page:

1. Entering a web page's URL into the address bar of the browser

2. Submitting a query to a search engine (visiting the result page)

3. Clicking on a hyperlink within a page

The first two actions are the possible first steps, when there are no other web pages in the web browser. These actions cause the system to assign that web page to an untitled thread. Peter starts with submitting the query "American car vs. Japanese" to the Google search engine and that will cause a node to appear in the visualization. As Peter did not choose any thread to which he could assign the web page, the results' page is assigned to an untitled thread (Figure 9). Threads' properties such as name, colour, and their tags can be further changed by choosing "Edit Thread" option from the menu that appears on the double click on the thread's root node (Figure 9). Peter decided to dedicate this thread to comparing companies from different countries and hence changed the thread's name to "compare countries" (Figure 10).

Peter has three options for visiting a web page whose hyperlink is within the current page, which is the result page for now. Explore-It has provided three options on the menu appearing on the double click on a hyperlink for opening a web page:

- Open in Current Thread

- Open in a Previous Thread

- Open in New Thread

Figure 9: Peter searched for "american cars vs. japanese cars" and the result page is automatically assigned to an untitled thread; by double clicking on the thread's root node, Peter can change thread's properties such as its names

Upon choosing either of the above options, the link will be opened in a new tab in the current browser window. The difference between these three options is the threads to which they will be assigned in the visualization. The mapping between the user's actions to the system's actions for these three user's actions is summarized in (Table 2).

User's Actions	System's Actions
1. Click on link	Opens the web page in the **current tab** and assign the page to the thread to which the current page belongs
2. Open in Current Thread	Opens the web page in a **new tab** and assign the page to the thread to which the current page belongs
3. Open in New Thread	Opens the web page in a **new tab** and assign the page to a new thread
4. Open in a Previous Thread	Opens the web page in a **new tab** and assign the page to one of the previous threads by choosing the thread
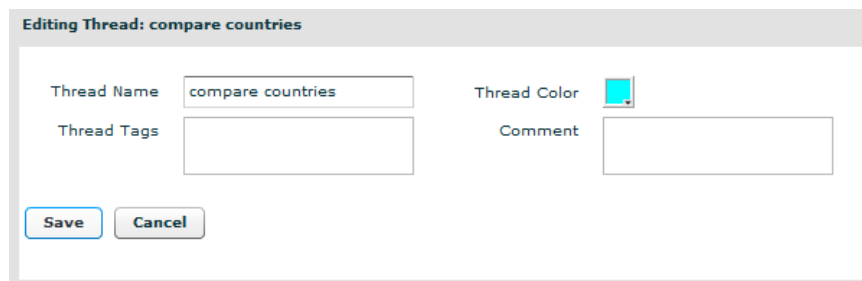
Figure 10: Editing thread popup panel that appears on choosing "Edit Thread" from double clicking on the thread's root node; Peter changed the default name of the thread "untitled1" to compare countries

By taking the first and the second actions in (Table 2), the web page will be assigned to the thread to which the current web page, containing the selected hyperlink, has been assigned. As the first hyperlink seems to be comparing

American cars with other cars, he chose to open that in the "compare countries"

thread, which was the current thread. Therefore, he chose "open in the current

thread" from the right click menu on the hyperlink (Figure 11). As Peter has

reached to the new web page from the previous page, there will be a link

between two nodes in the visualization representing the previous and the new

page. The node representing the new page appears below the node representing

its parent page (Figure 12).
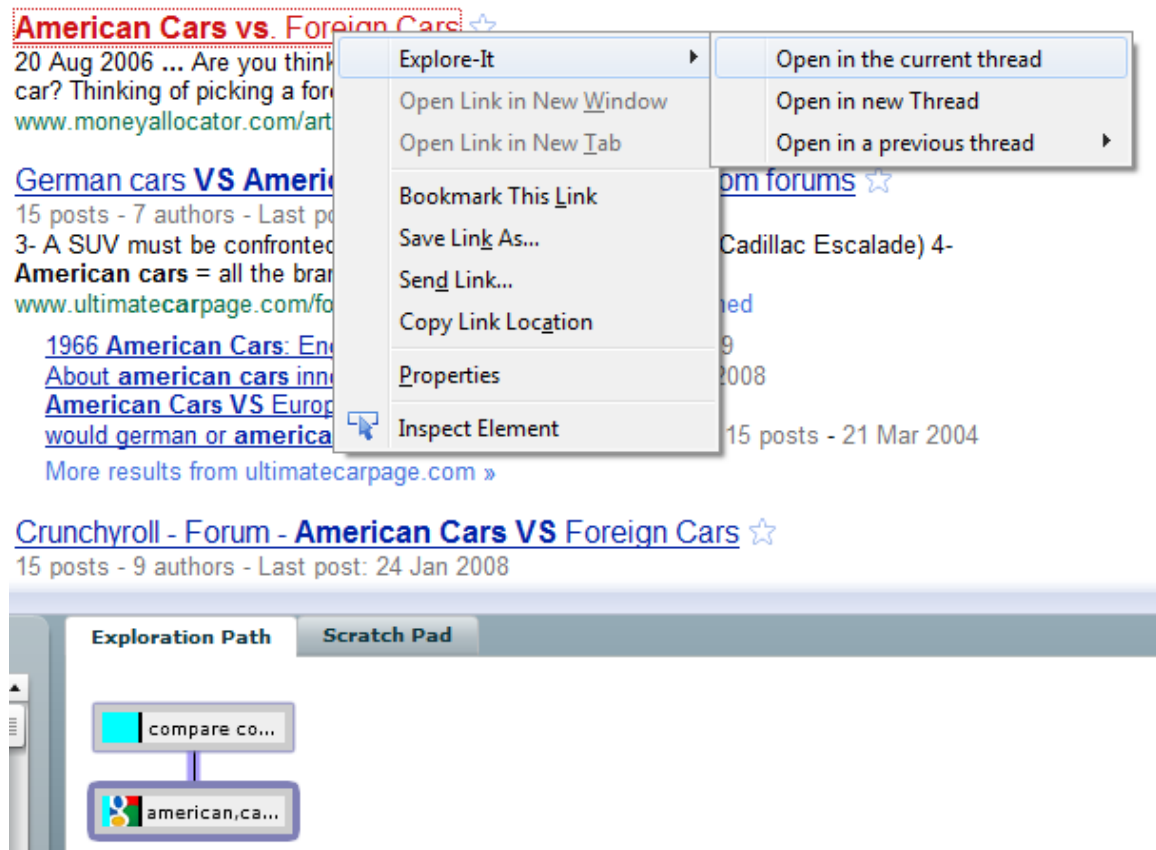
Figure 11: Options in opening a hyperlink; choosing the thread to which the new web page should be assigned
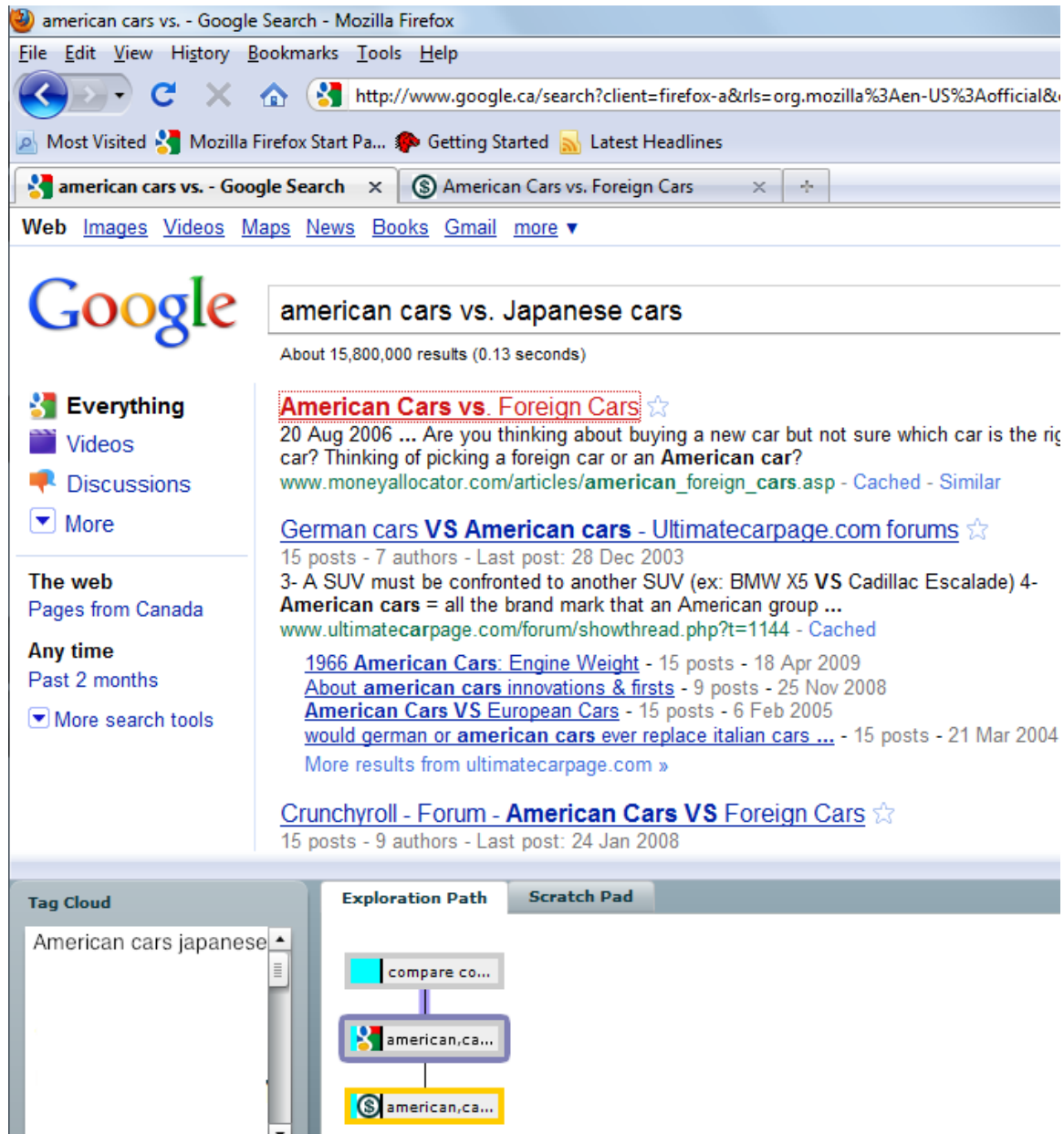
Figure 12: The first hyperlink in the result's page was chosen to be opened in a new tab and in the current thread; its corresponding node appears as a child of the results' page and an orange border of the node shows that the user has not selected the tab to visit the web page.

While visiting the recently opened web page, Peter found some useful

information such as the names of the web sites providing statics and surveys on

vehicles' performance and reliability. So, he decided to open those websites to refer to them when needed. As these new websites can serve a different goal from the current one, which was comparing cars from different countries, Peter chose to open those websites in a new thread. For opening the first website, he needs to create a thread. By choosing "Open in new thread" from the right click menu on the website's hyperlink (Figure 13), a popup panel appears for getting the basic information of the new thread (Figure 14).
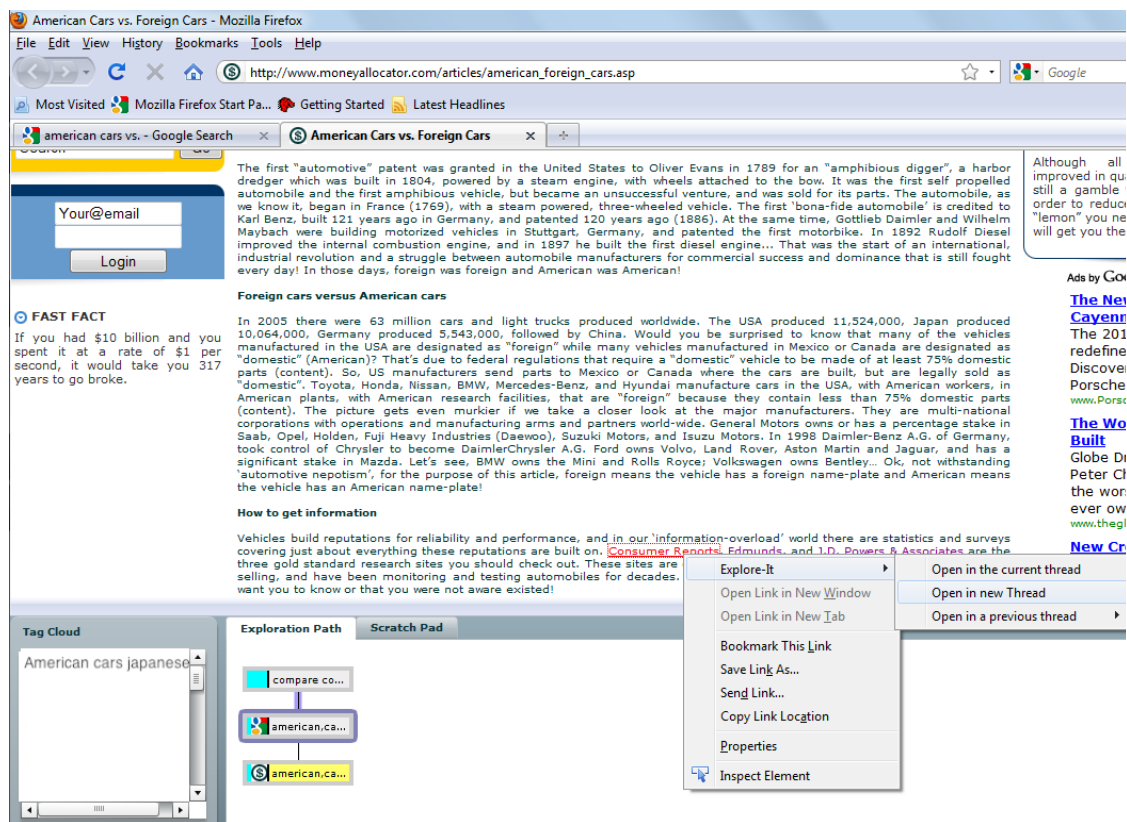
Figure 13: opening a web page in a new thread

How to get information

Vehicles build reputations for reliability and performance, and in our 'information-overload' world there are statistics and surveys covering just about everything these reputations are built on. Consumer Reports, Edmunds, and J.D. Powers & Associates are the three gold standard research sites you should check out. These sites are objective, in the business of testing and reporting, not selling, and have been monitoring and testing automobiles for decades. You can find information that manufacturers may not want you to know or that you were not aware existed!

Exploration Path   **Scratch Pad**

compare co...

american,ca...

american,ca...

Thread Name:
statistics surveys
Tags, comma seperated:
research, testing
Please assign a color to this thread:

Create

Cancel

Figure 14: Popup panel for creating a new thread; Peter has entered "statistics surveys" as the name of the thread and he also entered two tags "research" and "testing"

There are two other websites that Peter wants to assign them in the "statistics surveys" thread. Therefore, for the other two websites, he chose "Open in a previous thread" option from the right click menu (Figure 15). The system connects nodes representing those nodes to the selected thread's root node.
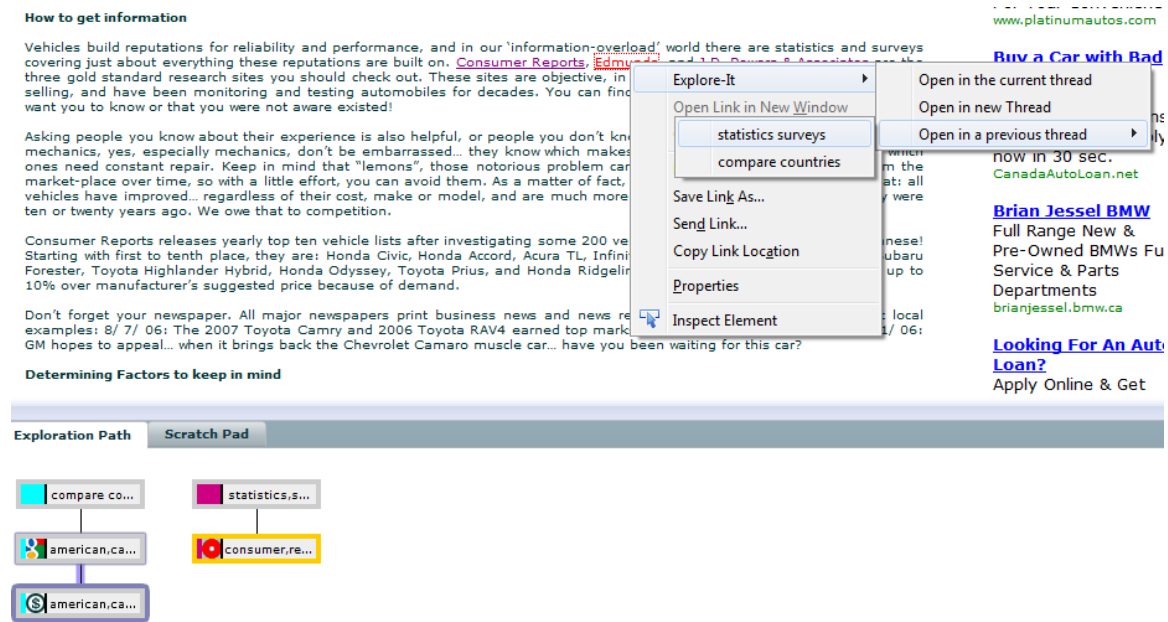
Figure 15: Opening a web page in one of the existing threads

After reading the current web page for a while, Peter realized that he didn't have enough information about the differences between different categories of the cars such as SUV, sedan, and minivan. Leaving the current thread without completing that, Peter starts a new search with the purpose of understanding the differences between these three different categories of cars. He names his new thread "suv, minivan, sedan". After visiting a couple of pages that have compared these three categories, he found that one of the important factors in comparing these classes is comfortability and he decided to dedicate a new thread to searching more about the comfortability of SUV and sedan (Figure 16). Soon, Peter realized that this new thread could be part of the "suv, minivan, sedan" thread and he moved the node representing the results' page of his first search

on comfortability of suv and sedan with all its children by choosing the option of

"Move subtree to another thread" from the double click menu (Figure 17).

Figure 16: Thread on the left is "suv, sedan, minivan" and the one on the right is "suv, sedan, comfortability"
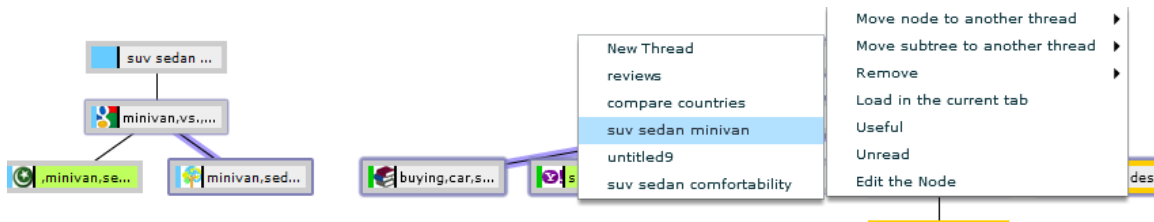
Figure 17: Moving a subtree from the thread "suv, sedan, comfortability" to the thread "suv, sedan, minivan"

The system moves the selected node with all its children to the selected

thread from the menu by removing the node from its current thread and

connecting that to the root node of the new thread (Figure 18).
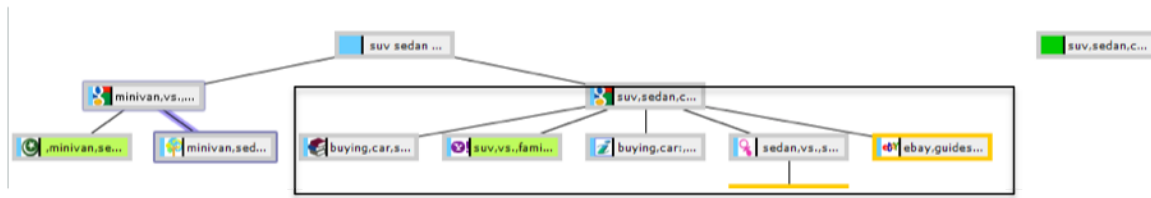
Figure 18: The node with all its children has been moved to the "suv, sedan, minivan" thread

As Peter did not need the thread "suv, sedan, comfortability" any more, he deleted that thread by choosing the option of "remove the thread" from the menu appearing on double clicking on the thread's root node.

5.5 Development and Implementation

The system is implemented as a Firefox Extension to accompany the user in her exploration of the Web. In a Firefox plugin, Javascript language is used for the logic part of the application and XUL for the interface part. XUL (XML User Interface Language) is Mozilla's XML-based language that provides building feature-rich cross platform applications that can run connected or disconnected from the Internet. An application with several elements shared among all the windows of a browser has an overlay file. The overlay determines the shared elements with their locations in the Firefox. Having Explore-It as part of each window/tab of the browser has several advantages comparing to a separate window design in which the system including the visualization part is just accessible in a standalone browser window. First, recording the exploration process and reflecting on the explored path does not interrupt the search process and does not increase the workload (Gersh et al., 2006) through moving to

70

another page or work with another system just for this purpose. Secondly, having the visualization of the user's trail during the exploration is beneficial for reflection for the same reason that chess players stare at the chessboard while thinking, instead of closing their eyes and thinking (Hutchins, 1995).

The interface part has been implemented in ActionScript 3.0 and XUL. Parts of the interface that required some changes to the main elements of Firefox such as adding a button to the status bar, changing the context menu, and overlaying a window on all the windows are implemented in XUL, while the parts including visualization of the user's trail and the tag cloud are implemented in ActionScript 3.0 (Figure 19). The backend is also written in Java and is hosted by Google App Engine (GAE). GAE runs web application on Google infrastructure and provide a distributed data storage service that features a query engine and transactions. The web application that is written in Java gets the interaction data from the interface (the ActionScript part) using RMI method and stores them in the App Engine datastore using Java Persistence API (JPA) interfaces (Figure 20).
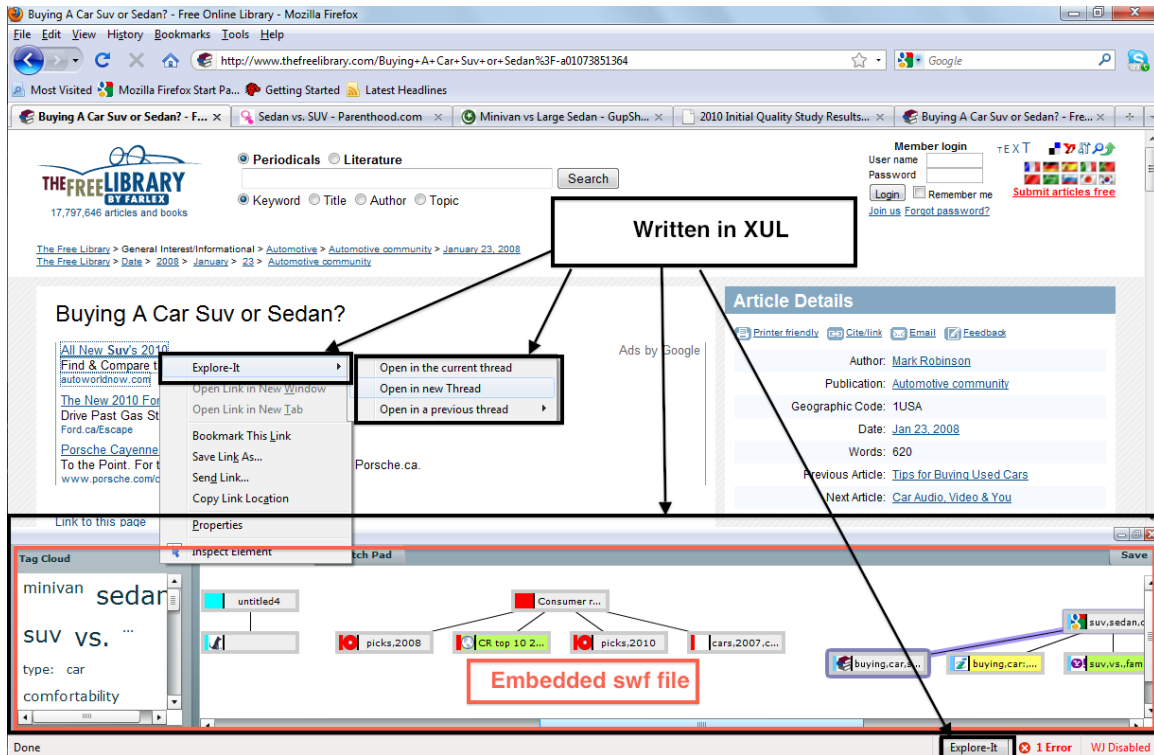
Figure 19: Parts of the interface that have been written in XUL are marked with black rectangles including the window at the bottom, button on the browser's status bar, and the changes made to the context menu of the browser; the red rectangle shows part of the interface that has been written in ActionScript 3 and is embedded in the XUL parts of the interface.
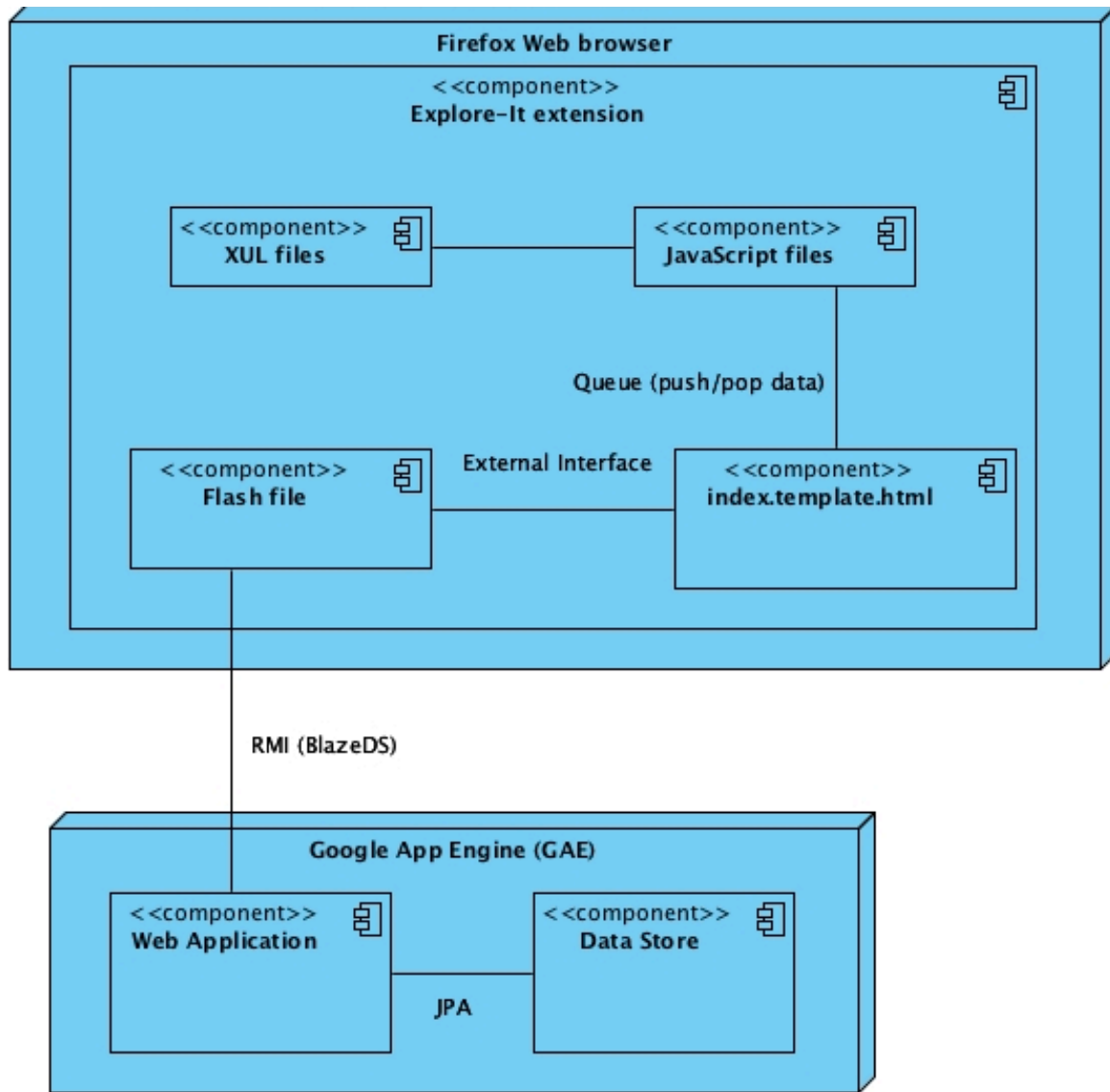
Figure 20: Deployment Diagram of Explore-It

6: METHODOLOGY

We conducted a case study 1) to understand the sensemaking processes that the users are involved in during exploratory Web search and 2) to understand how the proposed threaded model of exploratory web search can support such processes. In this study, we employed a mixed methods approach combining both qualitative and quantitative research methods. According to Tashakkori et al., a mixed methods approach is a type of "research design in which qualitative and quantitative approaches are used in type of questions, research methods, data collection and analysis procedures, and/or inferences" (Tashakkori & Teddlie, 1998). The purpose of using mixed methods research in this research is mostly collecting complementary data. We investigated whether the proposed model of exploratory search matched the user's activities in exploratory search by collecting and analyzing qualitative data such as users' opinions and the way users used the system. In addition to the qualitative data, we collected some quantitative data such as data derived from systems logs including the frequency of using each feature of the system and some subjective measures of usefulness or helpfulness of the system. The type of the mixed method is chosen based on the three criteria of timing, weighting, and mixing method. As in this research both the qualitative and quantitative data are collected concurrently, the timing of the mixed method is considered as concurrent. Although part of the qualitative data and part of the quantitative data

are collected in the case study and the rest are collected in the post-experiment questionnaire, it is still considered as concurrent because the first phase of data collection is not affecting the second phase. The second factor in planning a mixed methods procedure is the weight and the priority given to the qualitative or quantitative data. In this research, there is more emphasis on the qualitative study as the main goal of this research was to understand the behaviour of the users and see whether the threaded model of exploratory search can provide scaffolding for that process or not. Using the mixed methods notation for presenting the research strategy, our research methodology can be illustrated as **QUAL+quan.** This notion shows that the emphasis of the study is on the qualitative data and both data types are collected simultaneously.

Triangulation is one of the mixed method approaches that are usually used to check the results from different sources to ensure the validity of the results. According to O'Donoghue and Punch, triangulation is a "method of cross-checking data from multiple sources to search for regularities in the research data" (O'Donoghue & Punch, 2003). We employed triangulation of qualitative and quantitative data for some of the variables, which will be discussed in more detail in the quantitative analysis and qualitative analysis sections. The post-experiment questionnaire is designed to provide some qualitative data justifying and explaining the quantitative data from the system logs. For example, the frequency of usages of different features has been logged. The questionnaire has questions asking for the user's opinion on those system's features for finding the justification of the usages.

6.1 Case study

As a research method, case study has been employed in many situations to contribute to the knowledge of individual, groups, and organizations, social and political phenomena (Yin, 2008). "The essence of a case study, the central tendency among all types of case study, is that it tries to illuminate a *decision* or set of decisions: why they were taken, how they were implemented, and with what result" (Schramm, 1971). Although this definition focuses on decisions as cases in a case study, cases can be "individuals", "organizations", "processes", "programs", "institutions", and even "events". Likewise other research methodologies, case study also provides a systematic way of looking at events, collecting data, analyzing information, and reporting the results. Case studies can both generate and test hypotheses.

One of the determining factors in deciding about the research methodology is the form of the research questions. The main research questions in this thesis are "what" and "how" questions. Some types of the "what" questions are exploratory, such as "what are the sensemaking activities involved in exploratory Web search?". This type of question is a justifiable rationale for conducting an exploratory study with the goal of developing hypotheses for further studies. However, as an exploratory study, any of the research methods including case study, experiment, survey, and etc. can be used.

In contrast, "how" questions are explanatory and are more likely to favour case study. The "how" question in this study is "how does Explore-It's multi-threaded model of exploratory search provide support for the sensemaking

activities in the exploratory web search?". For approaching these exploratory

"what" and "how" questions in this thesis, we chose to conduct an *exploratory*

multiple case study to understand the user's exploratory search processes as

well as the sensemaking activities involved in the exploration process, in the

context. Controlled lab study usually separates the process from its context. As

we wanted the user to become immersed in the exploratory search process and

to care about learning and making sense of the information, case study was a

better approach to study the sensemaking activities and investigate whether

Explore-It is providing support for them or not. Also, time is another reason that

justifies the choice of evaluation in the wild comparing to controlled lab studies

for investigating exploratory web search. Performing a complete exploratory

search usually takes time and keeping the user in the lab for doing the search for

so long can make the user bored and consequently, the search will not lead to

exploratory search. According to Eraut, when time is extremely short, the

decisions have to be rapid. Therefore, the scope of reflection would become very

limited (Eraut, 1994).

6.2 Research Questions

6.2.1 Qualitative questions

There is one broad qualitative research question followed by a number of

qualitative sub-questions.

Broad Research question: *How do searchers use Exploration Threads in*

their exploratory search? (What is the user's conceptual model of the threads?)

Research question #1: What are the activities involved in the making sense of a topic while doing an exploratory search?

Research question #2: How does Explore-It support the users' activities in their exploratory search?

Research question #3: How do researchers describe their overall experiences using this system in their exploratory search process?

For answering this question, we conducted an interview with some of the participants after they worked with the system. They were asked some open-ended questions to describe their experiences. Further in our research, we included this data that we gathered from the interview in revising our solution to better fix the user's requirements in sensemaking.

Research question #4: How can Exploration Path (visualization of the user's trails) be improved to better support sensemaking?

Finally, we asked participants what they liked or disliked about Exploration Path and what it lacks to better support the kind of a search in which they were involved.

6.2.2 Quantitative questions and hypotheses

All of the quantitative research questions have been addressed through the questions asked in the interview with the users after they worked with the system.

Research question #1: *Does the structures in Exploration Path facilitate understanding and making sense of the information?*

The corresponding hypothesis is tested by Question #4 in the post-experiment interview:

"How do you rate the usefulness of the structures supported by Explore-It in Exploration Path for *understanding the information? (1=Not useful at all, 5=very useful)*"

Research question #2: *Does the structure in Exploration Path facilitate having more control over the exploration process?*

Question #5 in the post-experiment interview is designed to find the answer to this question:

"How do you rate the usefulness of the structures supported by Explore-It for having more control over the exploration process?" (1=Not useful at all, 5=highly useful)

Research question #3: *Is the mixed initiative approach useful in visualizing the user's exploration trail?*

This question is addressed by asking the question #8 of the interview from the user. (Appendix 1)

Research question #4: *Is the tag cloud component helpful for the user in tracking the changes in her interest during the search?*

Question # 10 asks for the user's opinion about the tag cloud:

"Tag cloud was helpful in understanding the changes in my search process (1=Strongly Disagree, 5=Strongly Agree)"

Research question #5: *Does Explore-It help the user in organizing the process and goal sequencing during the exploration?*

6.3 Task

6.3.1 Designing exploratory search task

Kules et al. provide a set of guidelines that should be considered in designing exploratory web search (Kules & Capra, 2009). Uncertainty, ambiguity, and discovery are the common aspects of exploratory search tasks. Users involved in an exploratory search tasks often feel lacking knowledge or contextual awareness in formulating queries and navigating complex information space. The exploratory search task requires browsing and exploration as sometimes all the available information are not indexed very well (White, Drucker, Marchionini, Hearst, & schraefel, 2007). Some of the common characteristics of exploratory search tasks identified in (Kules, Capra, Banta, & Sierra, 2009) include: 1. Uncertainty and ambiguity in information need, 2. Knowledge acquisition, comparison, or discovery, 3. A low level of specificity about information necessity and how to find the required information, 4. Provide enough imaginative context to enable the user to relate and apply the situation.

Kules et al suggested considering a delicate balance between broadness and specificity of the task. The tasks need to be broadly scoped to afford exploratory search behaviour. Meanwhile, there should be enough specificity in task descriptions so that the participants would be able to ground the task (Kules et

al., 2009). Considering all the above points, the following task's characteristics can well describe an exploratory search task:

- Low initial topic familiarity

- Require multiple items to be considered

- Some ambiguity even in the final answers

Below is an exploratory search task template suggested by Kules et al:

"Imagine you are taking a class called ---. For this class, you need to write a paper on the topic---. Find two possible topics for your paper. Find three books for each topic."

In designing the exploratory search task, it is beneficial to consider Borlund's recommendation on making the search topic interesting for the user (Borlund & Ingwersen, 2000). The challenge is to design interesting tasks for the participants that require the researcher to know the participants before designing appropriate tasks.

In this case study, participants were supposed to perform the following task:

"Imagine you are taking a class called <CLASS>. For this class, you need to write a paper on the topic <TOPIC>. Find one title for your paper and include a very rough outline for that. The rough outline can include the headings, keywords, and the title of references you might use in your writing"

The determining factors for finding the exact topic for each participant are the participant's interest to the topic and her knowledge of the topic. Therefore, the task template was supposed to be completed for each participant with different tasks according to their familiarity and their interests to the topic. However, when the first participant was given the task, he was willing to choose his topic himself as he had some general topic in mind for which he wanted to search. Therefore, the task template changed and instead of asking the participants to perform the above task, the participants were given the option of choosing a general area/topic for exploration. They were given two criteria to consider while choosing a general topic: 1. The topic should be interesting for them so that they are interested in and comfortable with spending time to learn about the topic, 2. The topic should be unfamiliar enough and they should not much background information on that. After each participant chose the topic, it was tested against the criteria of an exploratory search task. In this study, all the participants chose almost appropriate topic for performing a search on that. The final outcome of the task was a paper consisting of a title and a rough outline. The paper could be either an overview of the selected topic or they could focus on a specific facet of the topic for the paper. The details about the participants and the topics each selected are discussed in the next section.

6.4 Participants

In this study, 3 graduate students from the School of Interactive Arts and Technology at Simon Fraser University were recruited through email. A screening questionnaire was used to check the participants' search skills and

habits. All the participants were highly familiar with the Web search, including using variety of sources and search strategies. They all were familiar with Firefox and two of them reported using some Firefox extensions such as Zotero. They were asked about the way they were using Firefox regarding the use of tabs and windows. Two of the participants reported using multiple tabs in multiple windows, while the third participant reported using multiple tabs in one window. As predicted, they all had the same reason for changing the window; either the current window is too crowded to accommodate another page or the topic is getting different enough from the tabs in the current window so that they prefer to open a new window and continue browsing a different topic.

Participants were asked to find a general topic in which they are interested while it is an unfamiliar topic to them. The first participant (P1) wanted to buy a car and he needed to do a comprehensive search to find the right car that best matches his criteria. Therefore, his task was writing an outline for a paper about buying a car. The second participant (P2) stated that she wanted to learn about globalization in general. Her task was to explore this topic and come up with a title and a rough outline for a paper with the general topic of globalization. The third participant (P3) liked to know more about the 3D-TV. Therefore, his task was to come up with a title and a rough outline for a paper about 3D-TV. The first topics suggested by all the three participants were all appropriate enough to result in an exploratory search and there was no need for further discussions for finding another topic.

6.5 Data collection

The data, both qualitative and quantitative ones, have been collected in three forms including a pre-questionnaire, system logs, a post-questionnaire, and an interview. Explore-It was installed on the users' systems. Therefore, the task could be performed wherever the user could feel comfortable performing the search task; there is no restriction on where and when the user performs the search task. The details of the data collection regarding the data type, the collection methods, and the background of the method are in the following sections.

6.5.1 Questionnaire

There are three types of questions that can be used in a questionnaire: closed, open ended, and open response-options questions. In closed questions, the respondents have to choose an answer among the given answers as opposed to open-ended questions that no answer is suggested. Advantages and disadvantages of each question type are summarized in (Table 6). The questionnaire incorporates all the three types of questions because some forms are more appropriate for seeking particular form of response. However, according to some of the disadvantages of the closed and open-ended questions and the fact that open response-options questions can eliminate some of them, we assigned more weight to the open response-option questions i.e. the relative number of this question type to the other two is larger.

In designing the questionnaire, we considered answering to the questions in (Table 5 in Appendix 5) for each single question. Also, the questions were

written in a meaningful order in which the opening questions should be easy to answer, the questions should flow in some kind of a psychological order so that one leads easily and naturally to the next, and finally the question variety should be considered to avoid putting all the closed ones together or all the open-ended ones together at the end (Oppenheim, 1998). Also, considering that the participants become indifferent when they reach to the end, no important question was among the closing questions.

Pre-questionnaire: In order to assign an appropriate task to the participant, we need to know about the participant's interest to the topics and how familiar each topic is. In addition, we need to know the search and organizational skills and strategies of the participants. The questions of the pre-questionnaire are in the Appendix 1.

Post-questionnaire: The post questionnaire contains questions about the user's opinion about different features of the system. After investigating the system logs and the questionnaire, participants were scheduled for an interview for removing ambiguity from some of the data collected through the questionnaire and the system logs.

Interview: Face-to-face interview is the most useful type of data collection method, as the participants cannot be directly observed. Direct observation of participants might distract them from their exploratory search task, which requires them to be in the flow of task for making the experience as close to a real exploratory search experience. However, as researchers can bias the responses and not all the people are equally articulate and perceptive, there are some

85

limitations with this method that are considered in designing the questions. The purpose of this interview was collecting some qualitative data such as the user's mental model of the process and mental model of the system, which cannot easily be achieved just by looking at the system logs. In addition to note taking during the interview, the interview was audio-taped and transcribed for the further scrutiny.

6.5.2 System Logs

System logs is one of the unobtrusive method of collecting data. All the user's actions related to searching, organizing, and writing are logged by the system. These actions include creating, editing, and removing threads or nodes within the threads, writing comments, revisiting Web pages from Exploration Path and etc. These data were used to investigate how the users use the visualization in their exploratory search and particularly how they treated threads as an organizational structure for organizing their sensemaking activities. The system logs for the three participants are visualized by Tableu Software (Appendix 3).

6.6 Procedure

This study includes a pre-questionnaire, a 1-week case study, a post-questionnaire, and an interview if necessary. Each of the phases is explained in more details in the following sections.

6.6.1 Pre-experiment questionnaire

In the pre-questionnaire, participants were asked about their search skills and habits, their organizational habits during search, the tools that they have

used for organizing their search, and etc. The main purpose for asking these questions was knowing more about the participants' background to better explain the end results by taking the participant's background, skills, and habits into account while analyzing his/her behaviour using the system's logs and the post-questionnaire.

Delivering tutorials on how to use the system

After the interviews, participants were briefly trained on how to use the system and they were given 5-10 min to try the system and ask any questions that might be raised regarding how the system works.

The method by which the training materials were delivered to the participant is based on John Carroll's theory of minimalism in technical communication. He argues that training materials should be constructed as short task-oriented chunks as opposed to monolithic user manuals in which everything is explained in a long narrative form (J. M Carroll, 1997). According to his observation, users are already familiar with much of the described features described in the user manuals and only the information necessary for solving a particular task should be handy to them.

Therefore, the instructions were delivered in task-oriented chunks such as opening a web page, editing a node in the visualization, creating a thread, and etc.

6.6.2 Case study

Participants were given a week to complete their search task whenever

they had the time and energy to perform the task. They were asked to work with the system for at least 2 hours and inform us by email whenever they are done with the task.

The software is instrumented to log the participant's activities such as submitted queries, click stream, keystrokes, and etc. during the case study.

6.6.3 Post-experiment questionnaire

Once the participants were done with the task, they filled a post-questionnaire. Two of the participants were chosen for the interview about their experience with the system. The interviews were used for the qualitative analysis of the participants' behaviour. It also contained some quantitative questions aimed at testing the hypotheses mentioned in section 6.2. Participants were asked some open-ended questions for better understanding their search process and how Explore-It affected that.

6.7 Quantitative Analysis

The goal of the quantitative part of this study was twofold: 1. Obtaining statistical measures of the usefulness of the design and frequency of usages of various features of the system, 2. Obtaining complementary quantitative data to qualitative data. There are two major categories of quantitative variables: variables related to usefulness of the different design decisions and variables representing the frequency of feature usages. The variables of the first category are measured subjectively using the post-questionnaire, whereas the second category's variables are measured objectively through the system logs.

The variables in the second category are used in triangulation with the qualitative data from the questionnaire and the interview. Triangulation of data or using different research methodologies in the study of the same phenomenon helps in ensuring the validity of variables through cross verification from two different sources.

6.8 Limitations

One of the limitations of the study was the small number of participants which resulted in having just inferential statistics. Although the descriptive statistics gave us some insight of how the users generally used the system, inferential statistics are required to test the hypotheses and ensuring that we get the significant result.

Having considered the guidelines by (Borlund & Ingwersen, 2000; Kules & Capra, 2009) for designing exploratory search tasks, we let the participants choose their topic to ensure that they are interested in learning about the topic. Although we checked the topics chosen by the participants against those guidelines to ensure that they meet the criteria of being an exploratory search task, some of the searches did not result in a real exploratory search experience. One possible reason for that is the extent of user's involvement, which we thought that lack of knowledge of a domain and being interested to learn about that can guarantee a high level of involvement that apparently it did not.

6.9   Results

This section describes the results of the case study. First, the qualitative questions are addressed, and then the quantitative questions will be discussed using data from the system logs and the questionnaires. The qualitative data from these two sources and the interview gave us an understanding of how the multi-threaded model of exploratory search is supporting the sensemaking activities. For interpreting the results, we triangulated the usage-log data and the data from questionnaire; this helped us in assuring the validity of our interpretations of users' behaviour.

RQ: How do searchers use the threads in their exploratory search? (What is the user's conceptual model of the threads?)

One of the reasons for conducting a case study was to know more about the user's conceptual model of the Exploration Threads. According to the interview, users used the threads for organizing their exploration process without thinking about the final task, which was writing a rough outline for a paper. They used the threads for separating different aspects of their information need. Also, one of the users stated that although her primary goal for rearranging the threads was organizing her search, the structure of the threads was giving her an image of the domain to have in mind. This image had helped him to better understand the relationship between different pieces of information and to have a better sense of the semantic structure of the information in the domain under exploration.

RQ1: What are the activities involved in making sense of a topic while doing an exploratory search?

The task that the participants were given was a topic comprehension task, which is basically a sensemaking activity that involves information seeking, filtering, categorizing, comparing, and synthesizing. Information seeking, filtering, categorizing, and writing the outline were the activities recognized in the users' activities in their exploratory search. Planning the next steps in exploration was one of the common activities in exploratory search about which we found out in our interviews with users. The third qualitative question addresses how Explore-It could support these activities in which the users were involved in their explorations.

RQ2: How does Explore-It support the users' activities in their exploratory search?

According to the users' opinions, which were collected from the questionnaire and the interview, Explore-It could provide support for some of these activities. The Exploration Threads have been successful in supporting the organizing activities such as organizing the path, removing some of the visited web pages from the path, moving nodes of the Exploration Path to different threads once the user has gained a better understanding of the information in previous pages. One of the users used scratch pad, the free-form writing panel beside the Exploration Path, for planning the next steps in her exploration. One of the users used the scratch pad to write the keywords she wanted to use in her future searches. This can be considered as a planning activity. Another user

reported looking at the outline for getting an overview of the exploration process in order to plan the next steps in the search. Although the outline was not directly related on Explore-It features, this report was valuable for us as we realized that even the type of sensemaking task the user is doing has some role in the exploration process as well as the sensemaking per se. Considering the benefits of some sensemaking tasks such as topic comprehension and writing a paper outline, they can be incorporated as intermediate tasks in the design of exploratory search support systems to contribute to the main exploratory search task.

6.9.1 Quantitative Results

6.9.1.1 Quantitative data from the questionnaire

Some of the questions of the post-questionnaire were geared toward obtaining subjective measures of the usefulness of the different system's functionalities. The results of those questions are summarized in (Table 3) and visualized in (Figure 21).

Table 3: Users' ratings on usefulness of each of the system's functionalities. (4=very useful, 3=somewhat useful, 2=not particularly useful, 1=not useful)

	P1	P2	P3
1. Threads' structures for having more control over the exploration process	4	4	3
2. Threads' structures for getting a sense of general topic under exploration	3	3	3
3. Exploration Path for revisiting Web pages (navigation through Exploration Path)	4	3	4
4. Comments/notes on Web pages	2	2	1
5. Tag cloud for tracking changes in user's interest	1	3	1
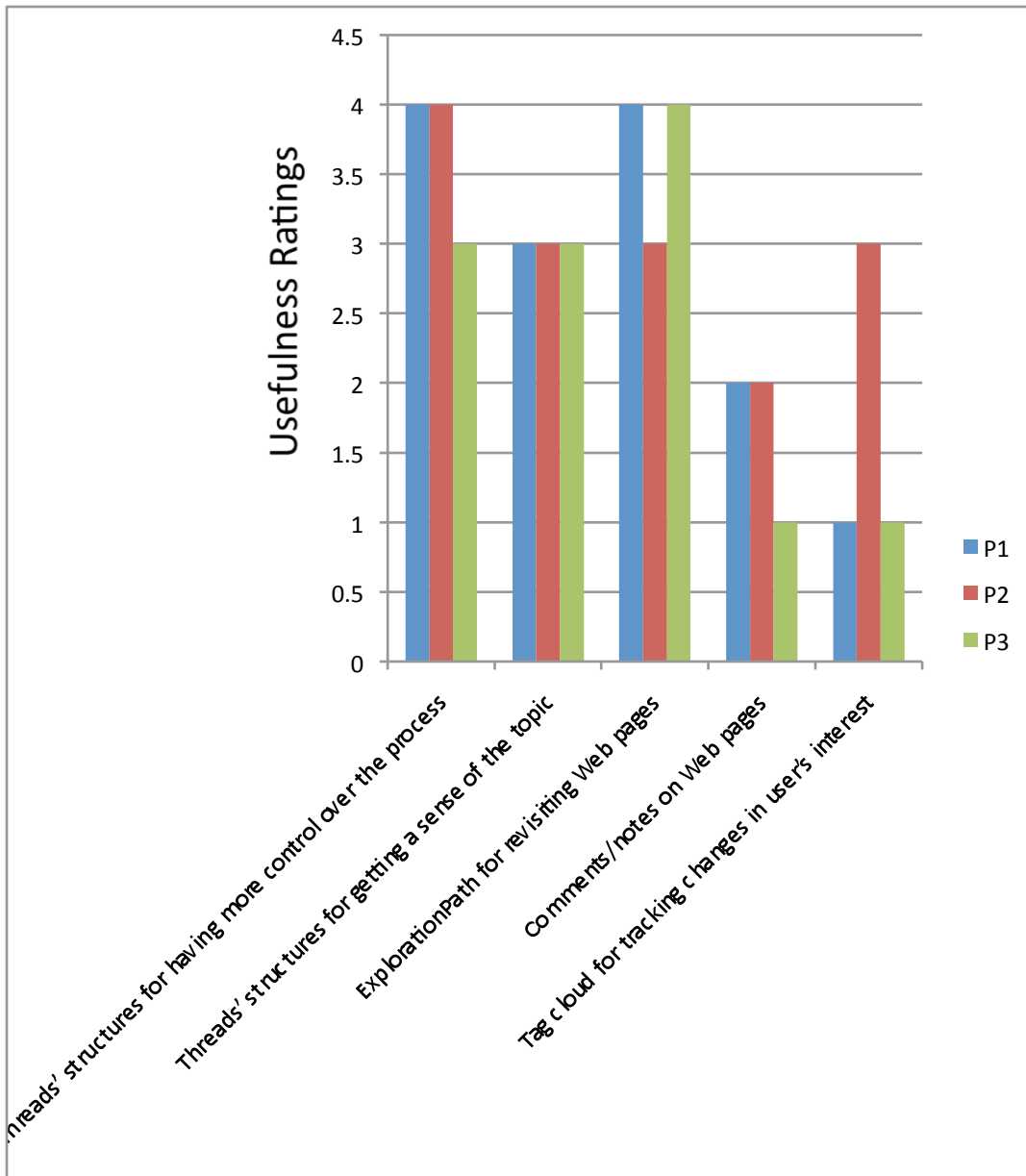
Figure 21: User's ratings on usefulness of the system's functionalities. (4=very useful, 3=somewhat useful, 2=not particularly useful, 1=not useful)

"Usefulness of the thread's structures for having more control over the search"

The first two participants have found the thread's structures useful for

having more control over the exploratory search comparing to not having

Explore-It. However, the third participant has found it "somewhat useful" which can be because of the fact that he visited much less number of web pages comparing to the other two participants (Figure 22). The advantages of the system in terms of giving more control to the user over the search process is more apparent when the user starts getting lost in the pile of web pages that she has opened. This usually happens as the number of visited web pages increases so that the user cannot keep all the pages actively in her mind to have the control of them.

"Usefulness of the thread's structures for getting a sense of the general topic"

All the users found the visualization "somewhat useful" for getting a sense of the general topic. Interestingly, the multi-threaded model was not equally useful for the sensemaking of the data as it was for the sensemaking of the process.

"Usefulness of the Exploration Path for revisiting web pages"

Explore-It provided the user with the option of navigating through the web pages from the visualization by selecting the "load in the current tab" option of the right click menu of the nodes in the visualization. Two of the users found this feature of the system useful whereas one of the users found that somewhat useful. When we asked for any reason supporting the "somewhat useful" answer, we found an interesting design idea suggested by the user. The user's statement is as following:

> *"The option of reloading the previously visited web pages from the visualization instead of having to find them among tons of open tabs is a powerful feature, however, the problem that I noticed with that was loading the page in the current tab as opposed to moving to the corresponding tab, if the web page is already open. Obviously, I prefer not to load my web browser window with multiple tabs of the same web page!"*
>
> *P2: The participant who searched for buying a car*

This complementary answer supports the quantitative findings of the overall usefulness of this feature, however, with some minor changes, it can be completely useful for the navigation purposes.

"*Usefulness of comments on web pages*"

None of the users found this feature useful. The users' opinion was also confirmed by the usage logs, according to which just one of the users used it once.

"*Usefulness of tag cloud for tracking changes in the user's interest*"

The largest discrepancy between the users' opinions appeared for this component of the system. Two of the users rated tag could as "not useful" for tracking changes in their interest. However, the third user found it useful for having a general overview of his search. He especially found the tag cloud useful for resuming the task in a new search session by being reminded of the keywords for which he searched in the last session. Interestingly, the tag cloud was found useful for facilitating task resumption, although facilitating task

resumption was not one of the goals in designing the tag cloud. Further

investigation on system logs and interview transcripts revealed some of the

factors that had caused the tag cloud not to be useful for tracking changes in

one's interest. First, one of the participants stated that whenever he noticed a

change in his interest, he created a thread to devote the new thread to that

interest. Therefore, at the moment of a change in interest, the user is completely

aware of the change and does create a thread, which removes the need for

keeping track of the changes in one's interest as it can be recognized by the

different threads. This actually revealed the close relationship between the user's

goal and the user's interest that was not apparent early in the design of the

threads. Secondly, the high similarity between the search topics and the papers'

titles in all three cases demonstrates lack of initial broadness of the search that

was one of the characteristics of exploratory search in its early stage (Table 4).

The narrower the search, the less likely it is for the user to face any changes in

her interest.

Table 4: Search topics and the paper's titles for all the participants

	Search Topic	**Paper's Title**
P1	Globalization	"Globalization: Yes or No?"
P2	Buying a car	"Finding the Right Car"
P3	3D TV	"Toward a True 3D Experience on Television"

As we believe that the tag cloud might be helpful in broader and more evolving exploratory searches, we decided to make it optional for the user to have this component beside the visualization or not. In fact, removing this component from the display can save some space for the threads in the Exploration Path that can be useful whenever it is not helping in the search.

6.9.1.2 Quantitative data from the usage-logs

For complementing the data collected from the questionnaire, we monitored the activity level of the participants to have an idea of to what extent each participant was committed to performing the task as well as to what extent they were involved in their search task. We divided the users' activities into two categories of search related activities and sensemaking activities. All the users' activities in which the threads and the nodes in the visualization were somehow involved are considered as sensemaking activities, as they are indicating a change in the user's understanding. These activities include renaming a thread, adding tags to a thread, moving a node from a thread to another one, removing a node, writing notes for the nodes, and etc. The following tables summarize the number of times that each user has performed each type of these activities (Figure 22, Figure 23, Figure 24, Figure 25).

		User		
Event		P1	P2	P3
Query submitted		14	35	13
Page opened in a new thread		17	29	4
Page opened in a previous thread		3	3	5
Page opened in the current thread		11	47	8
Grand Total		45	114	30

Number of visited pages =

Figure 22: Search-related activities of all the participants

	User		
Event	P1	P2	P3
Navigation from the visualiza..	7	3	6
Node changed its thread	6	1	4
Comment/Tags updated	22	2	1
Node is marked as useful		2	
Node removed	4	10	
Subtree changed its thread	2	4	
Thread deleted	11	15	1
Thread's color is edited:	5	6	
Thread's name is edited	5	6	
Grand Total	**62**	**49**	**12**

Figure 23: Sensemaking related activities for all the participants

	User		
Activity	P1	P2	P3
Search-related activities	45	114	30
Sensemaking-related activiti..	62	49	12
Grand Total	**107**	**163**	**42**

Figure 24: All the participants' activities
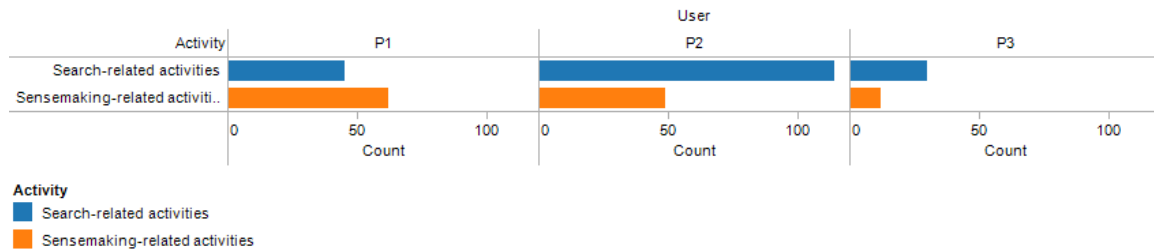
Figure 25: Diagram of all the participants' activities

According to the system logs, P1 has performed more sensemaking related activities than the search activities (65 comparing to 45). However, the other two performed more search related activities, i.e. spent most of their time exploring and searching than doing some sensemaking-related activities. P3, with the total number of 42 activities was generally less active comparing to the other two participants with 107 and 164 total activities.

6.9.2 Qualitative Results

In this section, we will review some of the users' opinions stated in the post-questionnaire or the interview that are not reported in the previous sections. In the post-questionnaire, participants were asked to write about their experience with Explore-It and describe how it helped them in their search. The following passage is P1's opinion about Explore-It:

*"Working with Explore-It for a while, I had the pattern of pages (the hierarchy, categories and colours) in my head so I could easily find the page that I visited before, and also **made me try to think of the category (thread) for each new***

*piece of information (webpage) that I encountered; this **helped me in more***

active reading/processing of information during my search**. Also, **Explore-It

saves the whole path automatically and give me the option to reorganize

***the path**, I don't like to manually bookmark each new page. I assume the path*

would be useful if I come back after a while (one or two years later) to remember

the search and possibly add to it."

P1: The participant who searched about Globalization

There are some interesting points to which P1 has pointed out in the above passage:

1. "made me think of the thread name for each new web page and this helped me in more active reading/processing of information in my search". This sentence confirmed the design decision of asking the user to assign each new web page to a thread for the reflection purposes.

2. "Explore-It saves the whole path automatically and gives me the option to reorganize the path". According to this sentence, the user seems to like the mixed initiative approach to constructing the visualization. This is confirming the design decision of making the construction of the visualization mixed initiative, so that the system simply visualizes the user's trail while the user can apply her own structure according to her understanding of the area.

We also explicitly asked whether they were satisfied about the level of their involvement in the activities related to construction of the visualization such as

making decisions about the thread to which the new web page should be assigned and choosing name for the new threads. They all were satisfied with the level of their involvement in choosing a thread for a new web page. However, one of the users preferred the system has suggestions for the threads' names. In addition, the same user stated that the system should be responsible for assigning a different colour to threads every time that a new thread is created. According to his opinion, threads' colours were just useful for distinguishing the threads.

Regarding the restructuring of the threads by moving the nodes to different threads and etc., two of the participants requested to have even more control over the restructuring. The rearrangement options that they liked to have were the option of changing the position of the threads so that they could place similar threads beside each other and the option of moving nodes up and down in the hierarchy. These two requests indicate that the user prefers to externalize her own mental model of the domain under exploration and keep that to have more control over the process as opposed to having her trail as a history of which web pages he visited.

However, naming the threads was not a completely satisfactory activity for all the users. One of the users stated that he prefers the system suggest some relevant names according to the web page's title or event its content.

The interview gave us some interesting information including the reasons behind some of the users' actions that could not be obtained otherwise. Some of the interview questions have been designed to understand the rationale behind

the user's actions. Among the restructuring mechanisms, there is an option for moving a node, which is representing a web page, to a different thread from its current thread. However, one of the users appropriated the usage of the "move node to another thread" feature in an interesting way. This feature is primarily for restructuring the structures when the new data/information/or the user's mental model cannot fit the previous structure. The user used this feature for moving a sub-tree of a thread to upper levels in the thread for making that sub-tree salient in the thread. However, the same user requested another feature with the opposite functionality, with which he is able to decrease the level of a sub-tree in a thread. The way the user appropriated the feature of moving a sub-tree and the new feature that he requested revealed the user's mental model of the levels in a thread. The level of the web pages in a thread seems to be indicating its importance level. In other words, the user was treating a thread as a hierarchy of folders in which the nodes in the higher levels were more important to one of the users.

In the first version of Explore-It, there was a panel for writing the paper's outline that required the user to change the display from the visualization to the writing panel. However, the first user found the writing panel a bit distracting as it happened to him to forget to switch to the visualization and that resulted in getting disconnected from the overview of his activity for a while. Therefore, we decided to remove that panel from the extension and instead ask the user to have an open document while searching. This helped the user to have the overview of his path implicitly, even if she is not paying direct attention to that.

6.10 Discussion

Having the existing sensemaking model as a reference, this study focused on investigating how people are involved in the sensemaking process during their exploratory search and how Explore-It can support construction of sensemaking representation during the exploratory search.

The exploratory search model suggested in this thesis introduced the concept of Exploration Thread, which is supposed to be the external representation of the user's thought stream during the exploratory search. The concept originated from the definition of exploratory search, according to which the information need has multiple dimensions. The results of this study confirmed the proposed multi-threaded exploratory search model, as the users used the threads to separate different aspects of their information need. This finding is according to the usage logs and the interviews. The Exploration Threads were used for organizing the visited/to-be-visited web pages to have more control over the multi-threaded exploratory search.

According to the Russel's sensemaking model, the sensemaking representations are updated by the user whenever the user's understanding of the domain is updated and it does not match the existing representation. Observing the restructuring of the representations among the users' activities during the search indicated that Explore-It has been useful in the restructuring of the representations. However, not all the restructuring activities were supported by Explore-It. Although moving the web pages between threads, removing threads or nodes, and moving a sub-tree in the thread to the higher level in the

same thread were supported as rearrangement options, the users liked to have more control over the restructuring such as placing the similar threads beside to each other.

The Exploration Threads or the sensemaking representation when looking from the lens of a sensemaking model, were reflecting two types of information: first, the user's understanding of the information space which was reflected through categorization of web pages in different threads; second, the user's exploration process and the parallel sub-topics that the users were exploring in their search. Therefore, the Exploration Threads can be considered as a two-fold sensemaking representation, representing both sensemaking of the data and sensemaking of the process of exploration.

The data from the questionnaire are in agreement with some of our decisions about the details of the methodology in addition to the design decisions. According to the questionnaire, the users used to use their web history rarely—about once a year or at most once a month. This result confirmed that the web history was not an appropriate system to consider for comparing to Explore-It, as people do not usually use their web history actively in their search. One of the reasons that people might look at their web history is finding a web page that they had forgotten to bookmark, only when it is difficult to find that by searching, otherwise searching again is preferred to looking into the web history.

Although we had planned to investigate the impact of thread's structures on the main task, which was writing an outline for a paper, by measuring the similarity between the threads' names and the headings of the outline, we didn't notice any

similarity between them. In fact, we found that not only the threads' structures and the outline were different for all the users; they were used for completely different purposes: the threads for organizing the search process and the outline for summarizing what they learnt in addition to writing that as a final task.

7: EPILOGUE

This research identified lack of support tools for supporting exploratory information seeking. The identified problem is addressed through designing and building a tool based on the existing sensemaking models and cognitive architectures, while incorporating the design guidelines from theories in learning, distributed cognition theory, and reflection in action.

7.1 Conclusion

Exploratory information seeking involves sensemaking activities; designing tools to support such tasks has been challenging due to the ambiguity of the situation facing the user. In exploratory web search, the ambiguity of the *process* plus the inherently unorganized nature of the exploration process increases the difficulty of designing such tools, in part due to the lack of a clearly defined target task. Providing support for organizing the exploration process that has often been overlooked in the design of exploratory support tools has been our core design consideration.

We designed Explore-It as a Firefox extension aimed at supporting exploratory web search. Explore-It is composed of two major components: an Exploration Path, a visual representation of the user's trail in exploring the information space; and a tag cloud representing the evolution of the user's interest during the search to keep the user aware of the changes to her interests.

Each of these components has been designed for supporting some of the important processes, underlying the sensemaking process, including organization of the sensemaking task, reflection on the process, tracking the evolution in one's interest to a topic during exploration, and communicating the information need with others. For example, Explore-It is using a mixed initiative approach to organize the exploration process based on the user's goals at each step of the exploration. The other component of Explore-It, the tag cloud, is providing awareness of the exploration.

In sum, we developed a multi-threaded model of exploratory web search building upon previous cognitive architectures, sensemaking models, learning theories, reflection in action, and distributed cognition theory. The broad research question in this research was: "how users use the structured representation provided by Explore-It for sensemaking of the information space and sensemaking of the process of exploratory web search". The research question was investigated by conducting case studies with three participants using Explore-It for performing an exploratory search.

Findings revealed various ways in which users were using the concept of Exploration Thread in their search. According to the data from the usage logs, the interview, and the questionnaire, the Exploration Threads were the externalized sensemaking representations, representing both sensemaking of the data and sensemaking of the process of exploratory search, i.e. users used the threads both for categorizing the resources and for organizing their search process. These two ways of using the threads were reflected in restructuring of

the threads as the sensemaking representations. According to the questionnaire data, the tag cloud was generally not useful for the search. Further investigation on the system logs and interview transcripts revealed two factors causing the tag cloud not to be useful: First, the lack of initial broadness of the search that resulted in less changes in the users' interest, and secondly, the existence of the threads removed the need for tracking the changes in one's interest as the user used a different thread for a following a different interest and therefore, caused the threads representing the changes to the user's interest.

Based on these findings, the study argues that the information systems that are aimed at assisting sensemaking in exploratory search should provide support for construction of sensemaking representation, where these representations can serve both as organizing structures for the resources and organizing structures for the exploratory search process. In addition, the structures should give sufficient control to the user for restructuring purposes. As the user's information need is evolving and the user's understanding of the domain is updating during the exploratory search, the restructuring options should be as broad as possible so that meet the user's need in restructuring of the sensemaking representations.

7.2 Future Work

We developed the concept of Exploration Thread as a sensemaking representation in exploratory web search in general, regardless of the search topic. However, different types of exploratory search, depending on their topic, require specific types of sensemaking representations that better match the topic and can better serve the final task, which can be writing a report, deciding which

product to buy and etc. We are interested to study the exploratory web search behaviour of users for different types of tasks while exploring World Wide Web to better investigate whether our multi-threaded model of the user's goal memory structure matches all types of final tasks in exploratory search. Also, conducting user studies with a larger number of users for testing the hypotheses is the next step in this research.

Explore-It has employed mixed initiative approach in visualizing the user's exploration path. However, it is using a *fixed* mixed initiative approach during the user's exploration, meaning that the method and the level of the user's and system's involvement in the constructing the visualization is not changing during the exploration. We are interested to incorporate *adaptive mixed initiative approach*, in which the system can adapt its involvement in construction of the visualization based on the user's situation and advancement in the process.

In addition, the Exploration Path has been designed to be a foundation for collaborative exploratory search by enabling the searchers to share their Exploration Paths with their colleagues for getting navigational advices on their explorations. Because this map can represent the searcher's mental model of the domain, it can contribute to establishing common ground by providing a meta-representation to the collaborators and the communication will be more efficient with more common ground. Completing Explore-It to support the described collaborative processes and conducting studies to investigate the usefulness of this approach is another part of the future work for this thesis.

REFERENCE LIST

Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks:
    personal Web space construction and organization. In *CHI-CONFERENCE-* (pp.
    41–48).

Albers, M. J. (2005). *Communication of complex information*. Routledge.

Anderson, J. R., & Bower, G. H. (1980). *Human associative memory: A brief edition*.
    Lawrence Erlbaum.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Lawrence
    Erlbaum.

Anderson, R. C. (1984). Some reflections on the acquisition of knowledge. *Educational
    Researcher*, *13*(9), 5.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A
    cognitive view*. Holt Rinehart and Winston.

Ayers, E. Z., & Stasko, J. T. (1995). Using graphic history in browsing the World Wide
    Web.

Bannert, M. (2006). Effects of Reflection Prompts when Learning with Hypermedia.
    Journal of Educational Computing Research, *35*(4), 359-375. doi:Article

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online
    search interface. *Online Information Review*, *13*(5), 407 - 424.
    doi:10.1108/eb024320

Bhavnani, S. K. (2002). Domain-specific search strategies for the effective retrieval of

healthcare and shopping information. In *CHI '02 extended abstracts on Human factors in computing systems* (pp. 610-611). Minneapolis, Minnesota, USA: ACM. doi:10.1145/506443.506508

Blanc-Brude, T., & Scapin, D. L. (2007). What do people recall about their documents?: implications for desktop search tools. In *Proceedings of the 12th international conference on Intelligent user interfaces* (pp. 102-111). Honolulu, Hawaii, USA: ACM. Retrieved from

http://portal.acm.org.proxy.lib.sfu.ca/citation.cfm?id=1216295.1216319

Boland Jr, R. J., Tenkasi, R. V., & Te'eni, D. (1994a). Designing information technology to support distributed cognition. *Organization Science*, *5*(3), 456–475.

Boland Jr, R. J., Tenkasi, R. V., & Te'eni, D. (1994b). Designing information technology to support distributed cognition. *Organization Science*, *5*(3), 456–475.

Borlund, P., & Ingwersen, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, *56*(1), 71–90.

Bransford, J. D., & Stein, B. S. (1993). *The IDEAL problem solver*. Freeman New York.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. *Cognitive development*, *3*, 77–166.

Brusilovsky, P. (2004). Adaptive navigation support: From adaptive hypermedia to the adaptive web and beyond. *PsychNology Journal*, *2*(1), 7–23.

Carroll, J. M. (1997). Reconstructing minimalism. In *Proceedings of the 15th annual international conference on Computer documentation* (pp. 27–34).

Collaud, G., Dill, J., Jones, C. V., & Tan, P. (1995). A distorted-view approach to

assisting web navigation. *New Paradigms in Information Visualization and Manipulation*, 95–100.

Csikszentmihalyi, M. (1998). *Finding flow: The psychology of engagement with everyday life*. Basic Books.

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. *Qualitative research in information management*, 61–84.

Doemel, P. (1995). WebMap: a graphical hypertext navigation tool. *Computer Networks and ISDN Systems*, *28*(1-2), 85–97.

Eklund, J., Sawers, J., & Zeiliger, R. (1999). NESTOR Navigator: A tool for the collaborative construction of knowledge through constructive navigation. In *Proceedings of Ausweb'99, The Fifth Australian World Wide Web Conference*.

Eraut, M. (1994). *Developing professional knowledge and competence*. Routledge.

Fisher, B., Agelidis, M., Dill, J., Tan, P., Collaud, G., & Jones, C. (1997). CZWeb: Fish-eye views for visualizing the world-wide web. *ADVANCES IN HUMAN FACTORS ERGONOMICS*, *21*, 719–722.

Fisher, B., Agelidis, M., Dill, J., Tan, P., Collaud, G., & Jones, C. (1997). CZWeb: Fish-eye Views for Visualizing the World-Wide Web. *PROC. OF THE 7TH INT. CONF. ON HUMAN-COMPUTER INTERACTION (HCI INTERNATIONAL '97*, 719--722.

Frécon, E., & Smith, G. (1998). WEBPATH-a three dimensional Web history. In *IEEE Symposium on Information Visualization, 1998. Proceedings* (pp. 3–10).

Fu, W., & Pirolli, P. (2007). SNIF-ACT: a cognitive model of user navigation on the

world wide web. *Hum.-Comput. Interact.*, *22*(4), 355-412.

Furnas, G. W., & Russell, D. M. (2005). Making sense of sensemaking. In *CHI '05 extended abstracts on Human factors in computing systems* (pp. 2115-2116). Portland, OR, USA: ACM. doi:10.1145/1056808.1057113

Gersh, J., Lewis, B., Montemayor, J., Piatko, C., & Turner, R. (2006). Supporting insight-based information exploration in intelligence analysis. *Commun. ACM*, *49*(4), 63-68. doi:10.1145/1121949.1121984

Grabowski, B. L. (2004). Generative learning contributions to the design of instruction and learning. *Handbook of research on educational communications and technology*, *2*.

Green, T. M., Ribarsky, W., & Fisher, B. (2009). Building and applying a human cognition model for visual analytics. *Information Visualization*, *8*(1), 1-13.

Hightower, R. R., Ring, L. T., Helfman, J. I., Bederson, B. B., & Hollan, J. D. (1998). PadPrints: graphical multiscale Web histories. In *Proceedings of the 11th annual ACM symposium on User interface software and technology* (pp. 121-122). San Francisco, California, United States: ACM. Retrieved from http://portal.acm.org.proxy.lib.sfu.ca/citation.cfm?id=288392.288582&coll=Portal&dl=ACM&CFID=81805257&CFTOKEN=95619658

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.*, *7*(2), 174-196. doi:10.1145/353485.353487

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Kadivar, N., Chen, V., Dunsmuir, D., Lee, E., Qian, C., Dill, J., Shaw, C., et al. (2009).

Capturing and supporting the analysis process. In *2009 IEEE Symposium on Visual Analytics Science and Technology* (pp. 131-138). Presented at the 2009 IEEE Symposium on Visual Analytics Science and Technology (VAST), Atlantic City, NJ, USA. doi:10.1109/VAST.2009.5333020

Kang, R., & Fu, W. (2010). Exploratory information search by domain experts and novices. In *Proceeding of the 14th international conference on Intelligent user interfaces* (pp. 329-332). Hong Kong, China: ACM. doi:10.1145/1719970.1720023

Kellar, M., Watters, C., & Inkpen, K. M. (2007). An exploration of web-based monitoring: implications for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 377-386). San Jose, California, USA: ACM. doi:10.1145/1240624.1240686

Kirsh, D. (1995). The intelligent use of space. *Artificial intelligence*, *73*(1-2), 31–68.

Kules, B., & Capra, R. (2009). Designing exploratory search tasks for user studies of information seeking support systems. In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries* (pp. 419-420). Austin, TX, USA: ACM. doi:10.1145/1555400.1555492

Kules, B., Capra, R., Banta, M., & Sierra, T. (2009). What do exploratory searchers look at in a faceted search interface? In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries* (pp. 313-322). Austin, TX, USA: ACM. doi:10.1145/1555400.1555452

Lin, X., Hmelo, C., Kinzer, C. K., & Secules, T. J. (1999). Designing technology to support reflection. *Educational Technology Research and Development*, *47*(3),

43–62.

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Commun. ACM*, *49*(4), 41-46. doi:10.1145/1121949.1121979

Newell, A. (1994). *Unified theories of cognition*. Harvard Univ Pr.

Newell, R. (1992). Anxiety, accuracy and reflection: the limits of professional development. *Journal of Advanced Nursing*, *17*(11), 1326-1333. doi:10.1111/j.1365-2648.1992.tb01855.x

O'Donoghue, T. A., & Punch, K. (2003). *Qualitative educational research in action: doing and reflecting*. Routledge.

Oppenheim, A. N. (1998). *Questionnaire design, interviewing and attitude measurement*. Pinter Pub Ltd.

Pedersen, E. R., Gyllstrom, K., Gu, S., & Hong, P. J. (2010). Automatic generation of research trails in web history. In *Proceeding of the 14th international conference on Intelligent user interfaces* (pp. 369-372). Hong Kong, China: ACM. doi:10.1145/1719970.1720033

Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis* (Vol. 2005, pp. 2–4).

Qu, Y., & Furnas, G. W. (2005). Sources of structure in sensemaking. In *CHI '05 extended abstracts on Human factors in computing systems* (pp. 1989-1992). Portland, OR, USA: ACM. doi:10.1145/1056808.1057074

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. *Schooling and the acquisition of knowledge*, 99–135.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The cost structure of sensemaking. In *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems* (pp. 269-276). Amsterdam, The Netherlands: ACM.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*(1), 101–130.

Schön, D. A. (1983). *The reflective practitioner*. Basic Books New York.

schraefel, M. C., Zhu, Y., Modjeska, D., Wigdor, D., & Zhao, S. (2002). Hunter gatherer: interaction support for the creation and management of within-web-page collections. In *Proceedings of the 11th international conference on World Wide Web* (pp. 172-181). Honolulu, Hawaii, USA: ACM.

Schramm, W. (1971). Notes on Case Studies of Instructional Media Projects.

Shneiderman, B. (1997). Designing information-abundant web sites: issues and recommendations. *International Journal of Human Computer Studies*, *47*, 5–30.

Simon, H. A. (1973). The structure of ill structured problems. *Artificial intelligence*, *4*(3-4), 181–201.

Spear, N. E. (1978). *The processing of memories : forgetting and retention / Norman E. Spear*. Experimental psychology series. Hillsdale, N.J. : New York :: Lawrence Erlbaum Associates ; distributed by the Halsted Press Division of Wiley.

Tabard, A., Mackay, W., Roussel, N., & Letondal, C. (2007). PageLinker: integrating contextual bookmarks within a browser. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 337-346). San Jose, California, USA: ACM.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: combining qualitative and quantitative approaches*. Sage.

Teevan, J., Cutrell, E., Fisher, D., Drucker, S. M., Ramos, G., André, P., & Hu, C. (2009). Visual snippets: summarizing web pages for search and revisitation. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 2023-2032). Boston, MA, USA: ACM.

Turetken, O., & Sharda, R. (2007). Visualization of web spaces: state of the art and future directions. *SIGMIS Database*, *38*(3), 51-81.

Wexelblat, A., & Maes, P. (1999). Footprints: history-rich tools for information foraging. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit* (pp. 270-277). Pittsburgh, Pennsylvania, United States: ACM. Retrieved from http://portal.acm.org.proxy.lib.sfu.ca/citation.cfm?id=303060

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. In *CHI '07 extended abstracts on Human factors in computing systems* (pp. 2877-2880). San Jose, CA, USA: ACM. doi:10.1145/1240866.1241100

White, R. W., Dumais, S. T., & Teevan, J. (2009). Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining* (pp. 132-141). Barcelona, Spain: ACM. doi:10.1145/1498759.1498819

Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation.

J. Am. Soc. Inf. Sci. Technol., *55*(3), 246-258.

WITTROCK, M. (2010). Learning as a Generative Process. *Educational Psychologist*, *45*(1), 40. doi:10.1080/00461520903433554

Won, S. S., Jin, J., & Hong, J. I. (2009). Contextual web history: using visual and contextual cues to improve web browser history. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 1457-1466). Boston, MA, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.sfu.ca/citation.cfm?id=1518701.1518922&coll=Portal&dl=ACM&CFID=81805257&CFTOKEN=95619658

Woodruff, A., Faulring, A., Rosenholtz, R., Morrsion, J., & Pirolli, P. (2001). Using thumbnails to search the Web. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 198-205). Seattle, Washington, United States: ACM. doi:10.1145/365024.365098

Yin, R. (2008). *Case study research: Design and methods*. Sage Pubns.

Zeiliger, R. (1998). Supporting constructive navigation of Web space. In *Proceedings of the Workshop on Personalized and Solid Navigation in Information Space* (Vol. 4, p. 4).

Zeiliger, R., Belisle, C., & Cerratto, T. (1999). Implementing a constructivist approach to Web navigation support. In *Proceedings of the ED-MEDIA* (Vol. 99, pp. 403–40).

Zhang, P. (2010). SENSEMAKING: CONCEPTUAL CHANGES, COGNITIVE MECHANISMS, AND STRUCTURAL REPRESENTATIONS. A QUALITATIVE USER STUDY.

Zhang, S., Farooq, U., & Carroll, J. M. (2009). Enhancing information scent: identifying

and recommending quality tags. In *Proceedings of the ACM 2009 international*

conference on Supporting group work (pp. 1-10). Sanibel Island, Florida, USA:

ACM. doi:10.1145/1531674.1531676

Appendix 1: Post-experiment questionnaire

1. How do you rate Explore-It's ease of use?(1=very difficult, 5=very easy)

2. How do you rate the usefulness of the visualization for revisiting a Web page that you have already seen it? (Assessing the usefulness of the interactivity of the Exploration Path) (1=Not useful at all, 5=highly useful)

3. (if you use bookmarks) How do you compare revisiting Web pages in Explore-It and in using bookmarks?

4. How do you rate the usefulness of the structures supported by Explore-It in the Exploration Path for *understanding the information*? (1=Not useful at all, 5=highly useful)

Any reason supporting your answer:

5. How do you rate the usefulness of the structures supported by Explore-It in the Exploration Path for having more control over the exploration process? (1=Not useful at all, 5=highly useful)

Any reason supporting your answer:

6. How do you rate the usefulness of comments/notes on web pages? (1=Not useful at all, 5=highly useful)

Please explain your purposes in writing notes for Web pages (If you wrote any)

a) Writing a summary of the Web page so that when I return to that Web page,

b) Writing my opinion about the Web page

c) Extracting the useful materials and saving them as notes

d) (For reflecting on what I read from that Web page) writing gave me time to digest the information

e) Other purposes:

7. The visualization matches the way I was thinking about my exploration process (1=Strongly Disagree, 5=Strongly Agree)

8. Were you satisfied with the level of your involvement in organizing the Web pages you encountered? (Please determine your satisfaction specifically for each of the following activities and provide an explanation justifying your answer)

a) Deciding on whether to open a Web page in a new thread, in a previously created thread, or in the same thread to which the current Web page belongs.

b) Naming the threads

9. Did tag cloud help you with your task? How?

10. How do you rate the usefulness of tag cloud for tracking the changes in what you were looking for or in your interest?

11. Did scratch pad help you with your task? How?

Appendix 2: Pre-experiment questionnaire

Age:--------------------   Sex: Female   Male

1. Which of the following browsers do you mostly use?

   a) Firefox   b) Internet Explorer   c) Google Chrome   d) Safari

2. Are you familiar with the browser's extension? (If No, go to the question#5)

   Yes   No

3. Have you ever used any Web browser extension for facilitating the search process? (If No, you can go to the question#5)

   Yes   No

4. Please fill the following table for the browser extensions that you use to facilitate your browsing and searching experience:

Browser Extension	Browser Type	What do you use for?	What is the most useful feature?

5. How do you usually use Tabs and Windows in Web browsers?

   a) Multiple Tabs in one Window

   b) Multiple Tabs in multiple Windows

   c) Just do everything in a single Tab

   d) Multiple windows (one tab in each window)

   e) Any other usage type:

6. When do you use "open in new WINDOW" instead of "open in new TAB"?

   a) To continue the browsing in a different enough topic to the tabs in the current window

   b) When I feel that this window is getting crowded by many tabs and I want to start a cleaner workspace (Not necessarily due to a big change in topic)

   c) Start a task after an interruption

   d) Any other reason for your decision:

7. Do you use bookmarks? (if No, you can go to the question 9)

   Yes   No

8. Do you organize your bookmarks in folders (do you create multi-level folder hierarchy of bookmarks)? (If Yes, what is the maximum folder level in your hierarchy?!)

   Yes   No

   2   3   4   5   6   7   more than7

9. Which of the following can be among your purposes when using bookmark?

   a. For organizing your current search

   b. For using that in the near future

   c. For saving just as a useful resource

    d. For spending a bit more time on that, hoping that I can remember it in the future better

    e. Other reason (please describe)

10. How often do you use your Web browsing history?

   a. Daily   b. at least once a week   c. at least once a month   d. Yearly
   e. Never

11. How do you rate your Web search skills?

    a. Successful: I can find whatever I am looking for easily

    b. Semi-successful: It usually takes a while to find out the appropriate keywords for the query

    c. Not very successful: I usually consult others to find better keywords

    d. Any other descriptions:

12. Have you ever experienced getting lost in the browsing? (If yes, Which of the following situation can explain your experience better?)

    a. Forgetting the very first reason that you started your searching/browsing

    b. Not knowing where exactly you are in your browsing (are you close to what you were looking for or too far from it)

    c. Following several parallel search

13. Have you ever used any tools for organizing your searching/browsing

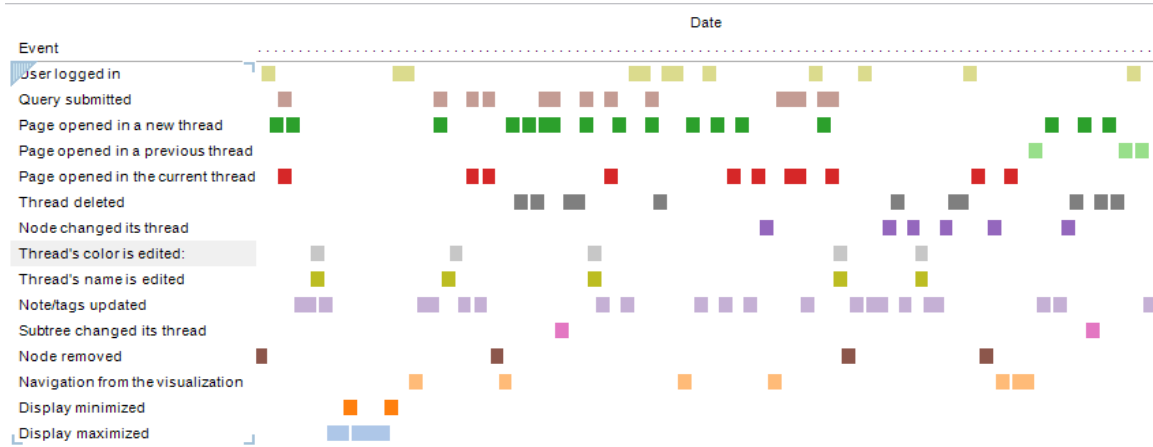process
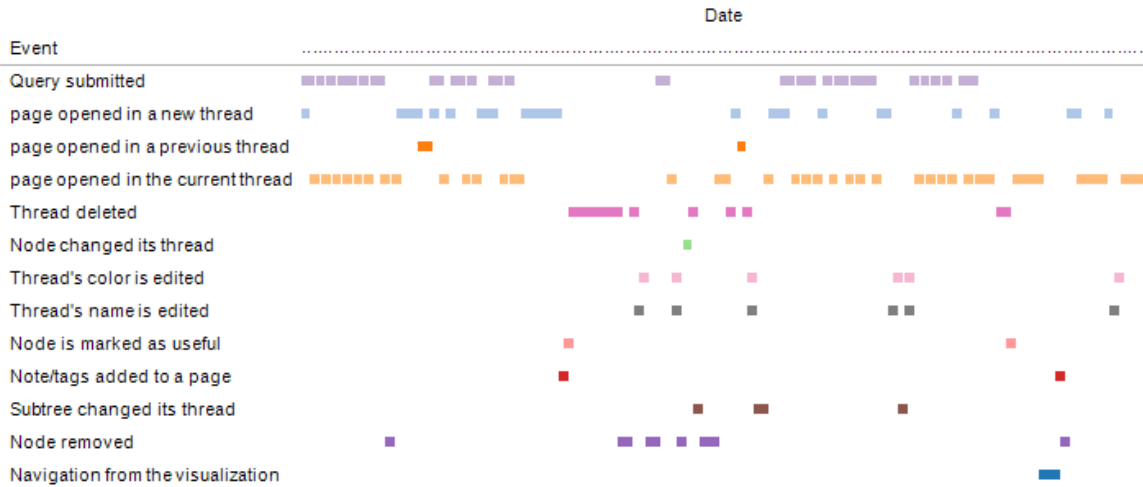
Appendix 3: System logs

Figure 26: P1's usage logs

Figure 27: P2's usage logs
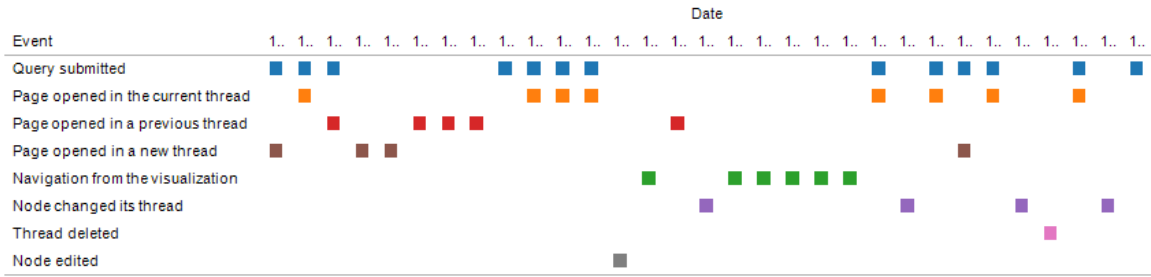
Figure 28: P3's usage logs

Appendix 4: Users' Exploration Threads and their papers' outlines
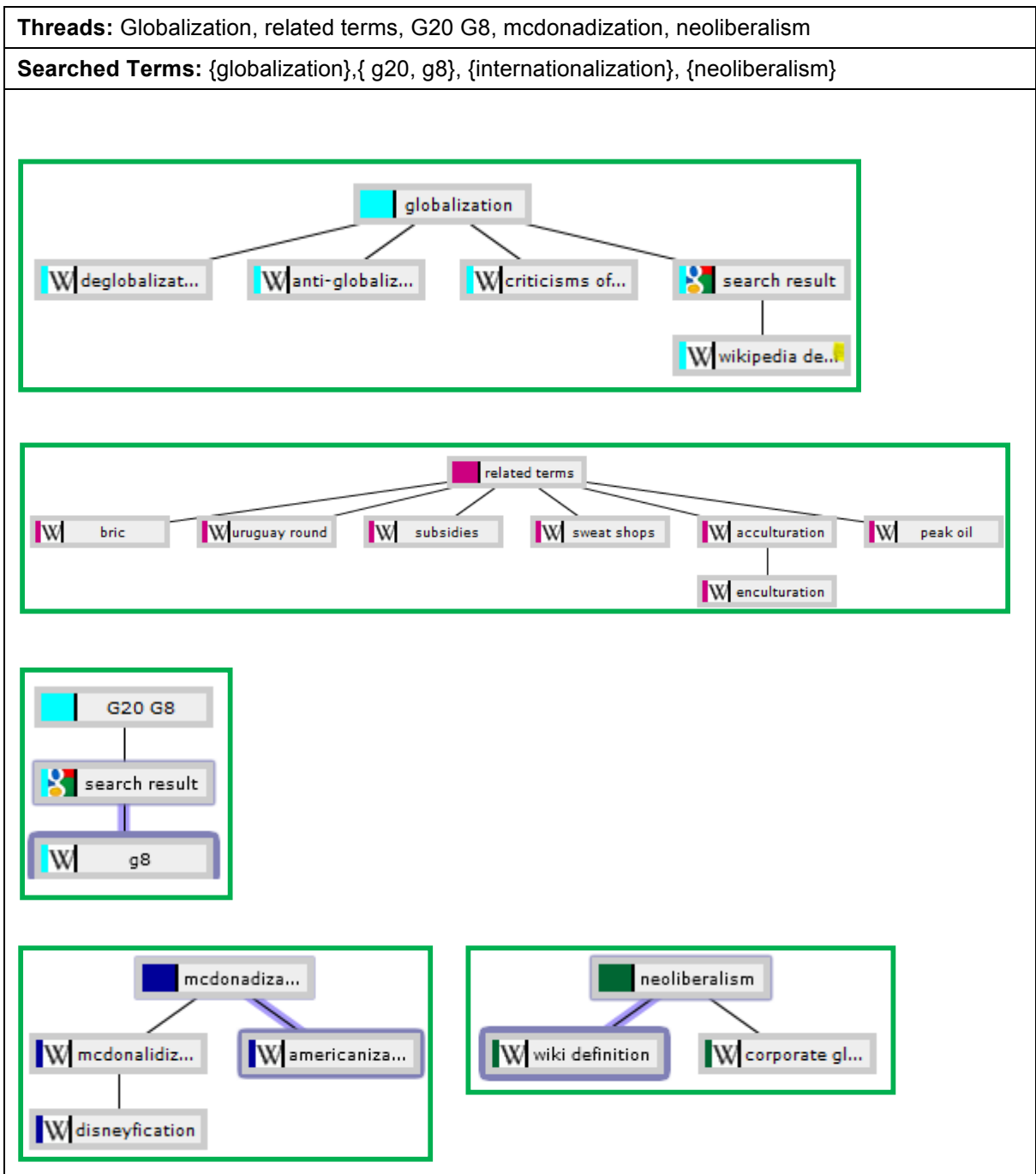
Threads: Globalization, related terms, G20 G8, mcdonadization, neoliberalism
Searched Terms: {globalization},{ g20, g8}, {internationalization}, {neoliberalism}

Figure 29: P1's Exploration Path

Globalization: Yes or No

Definition & Roots
Main dimensions
- **Economic globalization**
- **Cultural/social globalization**
- **Political globalization**

Globalization Measures

Effects

- **cultural**

- **environmental**

- **…**

Anti-globalization movement
- **The term**
- **Anti-corporate globalization**
    - **Neoliberalism**
    - **The ideology**
- **Anti-americanization**
    - **ideology**
- **Critisism**

Related organizations
- **Pro-globalization**

- **Anti-globalization**

Discussion

Figure 30: P1'a paper outline

Threads: compare countries, untitled3-4, reviews, consumer reports	
Searched Terms	{site:answers.yahoo.com american cars vs.}
	{consumer reports ten} {consumer reports ten 2009}
	{health pain suv}
	{pain comfortability suv}
	{site:answers.yahoo.com malibu}
	{vs. Leasing car}
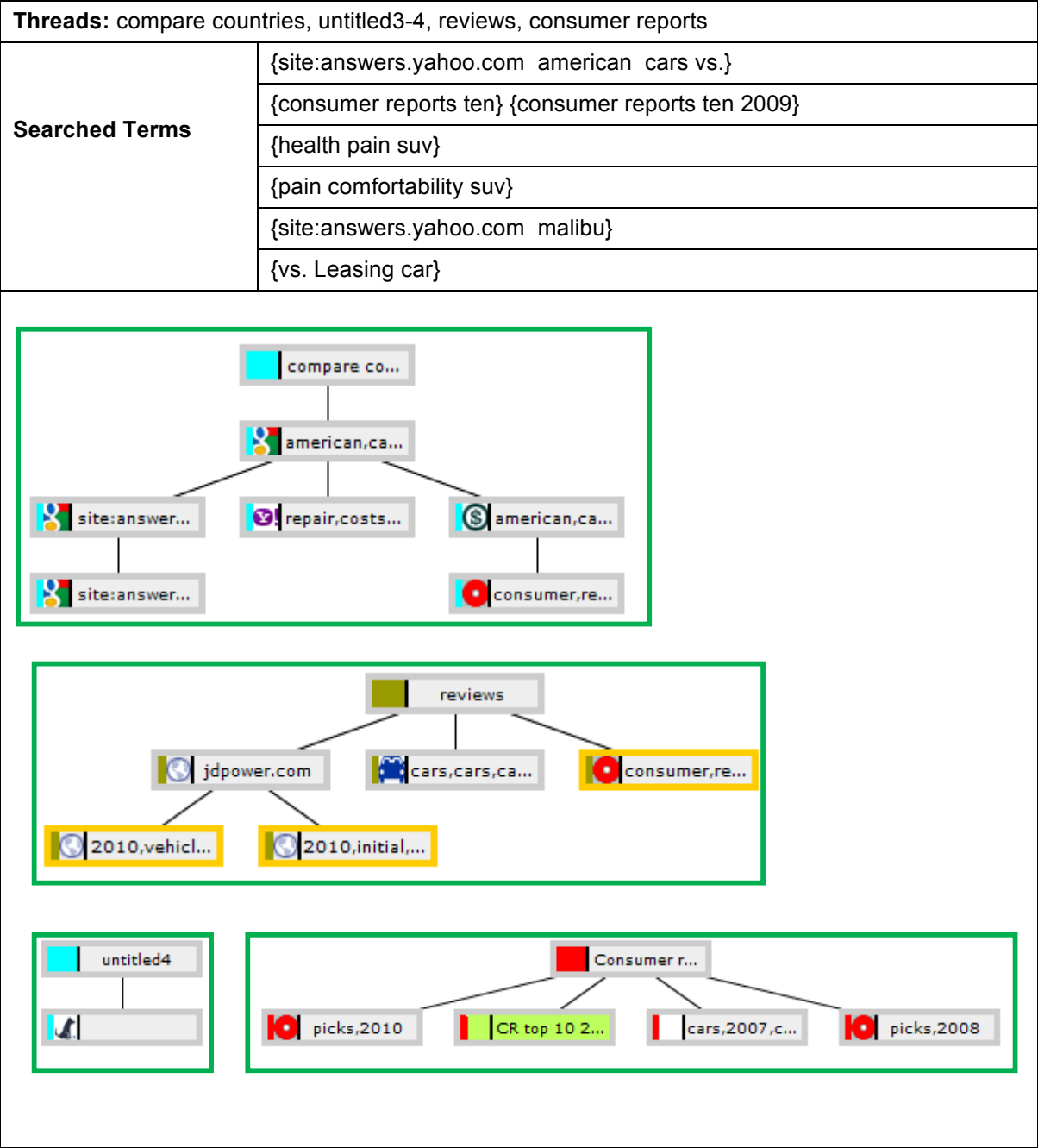
Figure 31: P2's Exploration Path

Finding the Right Car

What type: SUV vs. Sedan vs. mini-van
- **Kids**
- **Comfort**
- **Handling**
- **Fuel eff.**
- **Room**
- **Cargo**
- **Safety**
- **Look**

Which country?
- **Japan**
- **Korean**
- **European**
- **American**

Hybrid or not
- **Cost/benefit analysis and tax rebates**
- **Environmentally soundness**

Where to find reviews
- **JD powers**
- **Edmunds**
- **Consumer reports (just for top-10 lists unless you subscribe)**

New vs. used vs. lease
- **How many years to use**
- **Certified**

Dealing
- **What to do**
- **Auction**

Figure 32: P2's paper outline

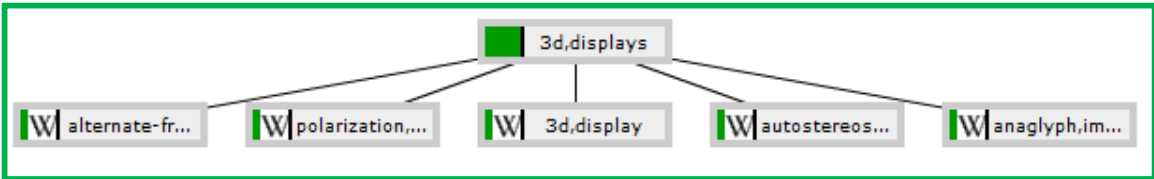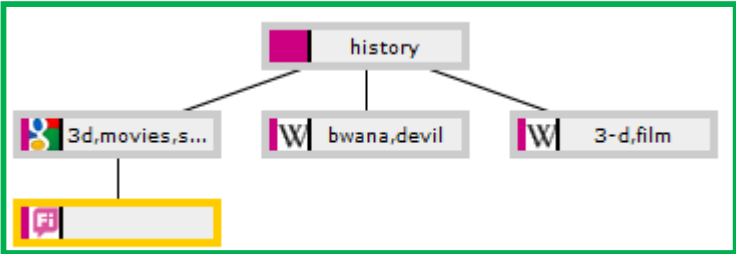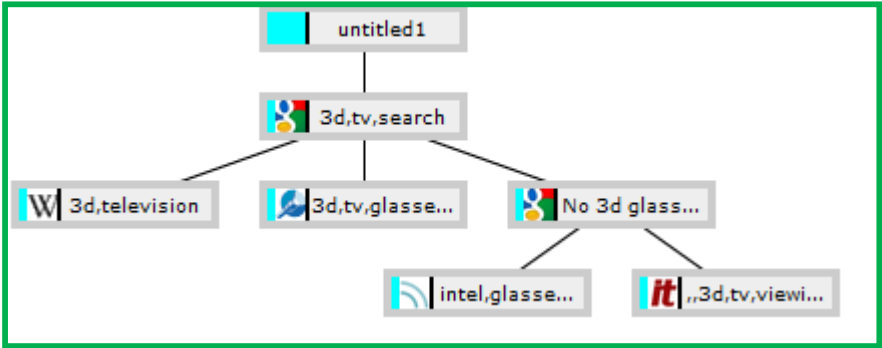| **Threads:** untitled1, history, 3d displays	
Searched Terms	{3d 2d tv success}
	{3d movies box office success}
	{3d movies success}
	{3d tv glasses}
	{3d tv}

Figure 33: P3's Exploration Path

Toward a True 3D Experience on Television

Abstract

Introduction

3d display techniques
- **Anaglyph image**
- **Polarization**
- **Alternate-frame sequencing**
- **Autostereoscopy**

Commercial trends
- **3D movies**
  - ○ **Avatar**
- **3D TV- State of Art**
- **3D TV- Best Practices**
- **3D TVs without glasses**

Limitations

Future Trends

Conclusion

Figure 34: P3's paper outline

Appendix 5: Designing the questionnaires

Table 5: Questions to ask in designing each single question of the questionnaire

1. Is this question sufficient in generating the required information?

2. Can the respondent answer the question correctly?

*An inability to response comes from three sources:

- Having never been exposed to the answer

- Forgetting (the questions are designed not to rely on the user's memory)

- An inability to articulate the answer

3. Are there any external event that might bias response to the question?

4. Do the words have the same meaning to all the participants?

5. Are any of the words loaded or leading in any way?

6. Are there any implied alternatives within the question?

7. Will the question be understood by the type of individual to be interviewed?

8. Is there any ambiguity in my question?

9. Are there any vague words or phrases?

10. Are there any questions too personal?

11. Do questions rely on feats of memory?

Table 6: Advantages and disadvantages of different types of questions in interview or questionnaire

Question type	Advantages	Disadvantages
Closed Questions	Provides the respondent with an easy method of indicating the answer (No need to think of how to articulate the answer)	Respondents are not allowed to give a different response to those suggested
	Prompts the respondent (No need to rely on memory for answering)	They suggests answers that respondents may not have considered before
	Responses can easily be classified which makes the analysis straightforward	
	Permits the respondent to specify the answer categories most suitable to their purposes	
Open-ended Questions	The respondent is allowed to answer in his own word (no influence by any specific alternative suggested by the researcher)	Respondents may find it difficult to articulate their responses
	Reveal issues that are most important to the respondent	Respondents may not give a full answer because they may forget to mention important points
	Respondents can qualify their answers or emphasise the strength of their opinions	Data collected is in the form of verbatim comments—it has to be coded and reduced to manageable categories (time consuming for the analysis)
		Respondents tend to answer open questions in different dimensions (for example about time and 'when' type of questions)
Open response-option questions	The researcher can avoid the potential problem of poor memory or poor articulation by being able to prompt the respondent into considering particular response options	It requires the researcher to have a good prior knowledge of the subject in order to generate realistic/likely response options before printing the questionnaire.
	Recording during the interview is relatively straightforward	