EXPLORING THE STRUCTURE OF TRAVEL PLANNING ON THE INTERNET

EXPLORING THE STRUCTURE OF TRAVEL PLANNING ON THE INTERNET

Bing Pan

Human Computer Interaction Group
Department of Communication
Information Science Program
Cornell University
Ithaca, New York
Email: bp58@cornell.edu
Mailing Address: 301 College Avenue, Information Science Building, Ithaca, NY 14850, USA
Telephone: 001-607-255-5530
Fax: 001-607-255-5196

Daniel R. Fesenmaier

National Laboratory for Tourism and *e*Commerce,
School of Tourism and Hospitality Management
Temple University
Email: drfez@temple.edu
Mailing Address: 201-cVivacqua Hall, 1700 N. Broad Street, Philadelphia PA 19122
Telephone: 215 – 204 - 5611
Fax: 215 – 204 - 1455

Submitted for publication in *Annals of Tourism Research*

May, 2003

BIOGRAPHICAL INFORMATION

Bing Pan is Post-doctorial Associate in Department of Communication and Information Science Program at Cornell University. His research areas include tourists' information search behavior, Human-Computer Interaction, cognitive information retrieval, and social aspects of digital libraries.

Daniel R Fesenmaier is a Professor in the School for Tourism and Hospitality Management and Director of the National Laboratory for Tourism and *e*Commerce, Temple University. His main research and teaching interests focus on the use of information and the Internet in travel decisions, the use of information technology for tourism marketing and the development of knowledge-based systems for tourism marketing organizations.

EXPLORING THE STRUCTURE OF TRAVEL PLANNING ON THE INTERNET

ABSTRACT

This research investigated the process of travel planning on the Internet based upon the assumption that travel planners navigates through the travel information space according to their idiosyncratic semantic mental models. Based on a travel planning exercise, this research shows that travelers have a variety of semantic mental models regarding a destination and they tend to search for information hubs on the Internet. Their information search process can be deconstructed into different episodes and chapters reflecting the specific problem being addressed. The theoretical contribution of this study and implications for designing better information technology for travel planning are discussed and future research directions are detailed.

Keywords: Travel planning, information search, Internet, semantic mental models

EXPLORING THE STRUCTURE OF TRAVEL PLANNING ON THE INTERNET

INTRODUCTION

Tourism is an information intensive industry in that tourism organizations rely on the exchange of information with travelers through various channels to market their products and build customer relationships (Poon 1993). Additionally, travelers depend on travel-related information to satisfy their various information needs including travel planning (Vogt and Fesenmaier 1998). However, travelers are often overwhelmed by huge amount of information and not able to locate the information they intend to find (Pan and Fesenmaier 2000). Recent studies of online tourism have confirmed that travel planning on the Web is a frustrating experience (Radosevich 1997; Stoltz 1999). One important cause for the usability problems of the Internet as a travel information source is that travel information search is highly dynamic and contingent on an information searcher's background knowledge, his or her individual characteristics, search tasks, and stages of travel planning (Jeng 1999). Furthermore, research has shown showed that people use a surprisingly large number of different words to describe the same concept, which makes it difficult to accommodate these different representations in an information system (Furnas, Landauer, Gomez and Dumais 1987). A variety of systems have been built on the Internet to facilitate travel information search. Numerous travel related web sites offer a wide range of services including online booking, virtual tourist communities, and search functions on destinations, accommodations, activities or attractions. Various computer agents have been developed to facilitate travel planning on the Internet (Linden, Hanks and Lesh 1997; Rich and Sidner 1998). In addition, intelligent recommendation systems using artificial intelligence and expert system techniques based on behavioral data have been designed (Ricci and Werthner 2001; Hwang and Fesenmaier 2001; Klicek 2001). However, the effectiveness of these systems remains unclear and have not been widely adopted.

Since the Internet possesses the potential for high levels of interactivity and can be customized to a great extent (Newhagen and Rafaeli 1996), the Internet should be able to provide tailored information to users according to their idiosyncratic preferences. Bishop and Starr (1996) argued that a thorough understanding of information search process is needed in order to design customized interfaces for information systems. Thus, without a thorough understanding of travel planning and travel information search behavior on the Internet, high-quality travel related information deliverance is not plausible. Based on the literature in tourism, Human-Computer Interaction (HCI), and information science, the goal of this research was to explore the structure of information search within the context of travel planning on the Internet. This study first examined relevant literature on travel information search, the Internet as a travel information source, navigation on the Internet and usability of the Internet. Second, protocol analysis and process tracing method were used to explore the information processing and decision behavior in the travel information search. Finally, the implications for designing useful information technology for online tourism were discussed.

TRAVEL PLANNING ON THE INTERNET

Travelers make decisions within an information-rich environment, which includes various travel information sources. Information search can be divided into internal search, which is a scan of long-term memory for relevant product knowledge, and external search, which happens when internal search can not provide sufficient information (Bettman 1979; Engel, Blackwell and

Miniard 1990). In terms of external information sources, travelers rely on both marketing-dominated and non-marketing dominated information sources to search for travel related information and plan their trips. The former information sources include advertising and commercials on mass media, travel brochures, guidebooks from clubs and welcome centers; the latter includes friends, relatives, and personal experience. Fodness and Murray (1999) explored travel information search strategies using clustering analysis, and distinguished seven search strategies: prepurchase mix, tourist bureau, personal experience, ongoing, on-site, automobile club and travel agency. They argued that these strategies can be arranged along three dimensions: temporal (ongoing search versus prepurchase search), spatial (internal search versus external search), and operational (contributory information source versus decisive source) dimensions.

A recent study by NFO Plog Research showed that the Internet has become one of the most important information sources for travel information acquisition; about 95 percent of web users used the Internet to gather travel relate information; about 93 percent indicated that they visited tourism web sites when planning for vacations; and nearly one half said they used email to gather travel related information (Lake 2001). The choice of different information sources has been studied by Ratchford, Talukdar, and Lee (2001); they argued that the choice of information sources is determined by the cost and benefit of accessing information through those sources. For example, while traveling, traveler may use welcome centers, sign posts, or mobile devices for *en-route* information search since they are convenient to use and demand relatively little cognitive effort. For detailed travel planning prior to the trip, the Internet is a better choice since it is universally accessible and contains a large quantity of information (Pan and Fesenmaier 2000).

Navigation on the Internet

The Internet is an enormous hypertext system which is composed of interconnected text and information nodes. Travel information searchers navigate through the travel information space in order to obtain relevant information. Accordingly, the value of the Internet is determined by the ease of access and navigation and is uniquely related to its hypertext nature and the information searcher's navigation behavior. Research in hypertext has shown that three methods can be used to carry out information search tasks in a hypertext system: (1) the user can traverse through a set of links to reach interested pages; (2) the user can navigate through the documents using a representation interface (such as using a graphic bookshelf to represent the organization of online books); and, (3) the user can query all documents through keywords search to locate relevant documents (Conklin 1987). The main advantage of hypertext is that large amounts of information can be accessed rather quickly and the organization of the information is flexible. On the other hand, hypertext does not have a conventional structure that leads the user through the networks of documents and the user is completely unrestricted in terms of where to go and which hyperlink to click; therefore, hypertext-based systems require users to exert more cognitive effort.

Two different theories have emerged to provide explanations for hypertext navigation: (1) information foraging theory (Pirolli and Card 1999) and, (2) a cognitive model of web design and navigation (Bollen 2001). In an analogy with the food foraging behavior of living organisms, information foraging theory is a general model describing how people use different strategies and technologies to search for information in response to the changing information environment. The theory posits that information searchers use proximal cues to identify important information for further exploration or consumption. The concept of information scent describes how information searchers identify valuable information from "snippets" of proximal cues (represented by link anchors on the Web). In an empirical study, the value of information scent was measured by vectors of words in the documents in relation to the information searcher's intention. On a web page the decision of which link to click depends on the assessment of the values of information scent of all the evaluated links (Chi, Pirolli, Chen and Pitkow 2001). Alternatively, Bollen (2001) proposed a cognitive model of web design and navigation whereby he argued that shared knowledge is necessary for hypertext navigation. He further argued that user expertise, hypertext network structure, navigation strategies, domain knowledge and mental models all contribute to the success of hypertext navigation. In his research, the mental models of the users are represented by their navigational paths and the model of the system is represented by the hyperlink structure of the web site. In explaining the navigation process, Kim and Hirtle (1995) argued that users of a hypertext system need to perform several tasks at the same time. Informational tasks, which is reading and understanding the contents presented in the linked nodes; navigational tasks, which is planning and searching through links; and also management and negotiation of the previous two. Failing the first task may lead to disorientation in hypertext systems. Similarly, according to Nakayama, Kato and Yamane (2000), web pages can be categorized into index pages and content pages, representing indexical information and content information; indexical information is used for navigation purposes and content information is used by users to process information in depth. Consequently, a dichotomy of information processing can be proposed to explain the navigation process in the hypertext system: (1) Navigation, whereby an information searcher browses through web pages in order to find more relevant web pages; and, (2) Reading, whereby the information searcher remains on certain web pages for a longer period of time in order to more fully process information contained on the page.

Cognitive Information Search and Retrieval Theories and Usability

According to Jacob and Shaw (1998) there are two paradigms which guide most information retrieval research. The physical paradigm is based on an analogy with mechanical systems which does not take into account the user's cognitive mode; on the other hand the cognitive perspective of information retrieval argues that "any processing of information, whether perceptual or symbolic, is mediated by a system of categories or concepts which, for the information-processing device, are a model of his world." (de Mey 1977, pp. xvi-xvii). The cognitive perspective of information retrieval views information as subjective instead of objective. That is, information only makes sense when it is assimilated into the mental model and knowledge structure of the information receiver. Accordingly, effective information retrieval depends on the congruence between the cognitive structure of an individual user and the knowledge representation of the information system (Shera 1965). Since each individual's idiosyncratic mental model is influenced by many socio-cognitive variables and is hard to capture and analyze, it is easier to explore the shared mental model and knowledge representation of a user group or knowledge domain (Allen 1996). In addition, since representation of knowledge is language-based, it is essential to investigate the role languages played in the communication process (Blair 1990). Similarly, Ingwersen (1992) argued that in information retrieval process, data in an information system have been transformed into information in relation to the mental model of the user and subsequently, the knowledge state of the user is changed. The concept of "polyrepresentation" of information can be applied to both the user's mental model and the system's information space. The linkage (or overlapping) between different representations of semantic models of the user's and the system's can reduce the uncertainty of information retrieval process by eliminating lexical ambiguity and providing contextual information. Furthermore, the interaction between a user and an information system is mediated by user interface. Users always need to generate queries to match the system's language (Beaulieu 2000), which represents compromises between his or her information needs and the need to adapt to the information system. In general, the cognitive information retrieval literature views the degree of congruence between the mental model of information users and the semantic model of the information systems as determining the usefulness and efficiency of the system.

A Conceptual Model of Travel Information Search on the Internet

Following cognitive information retrieval literature, travel information search on the Internet can be viewed as the interaction between information searchers and the information space (the part of the Internet related to tourism and travel destinations) in the context of travel planning and travel information search. Travel information space contains different types of information provided by different parties in tourism industry along with technology professionals in order to market tourism products and communicate with travelers. Based on this framework, a model of travel information search on the Internet can be proposed (Figure 1). Three components constitute the interaction: a travel information searcher, the interface, and the travel information space. Traveler's situational factors, knowledge and skills regarding traveling and the travel information space, contribute to effective travel information search; the travel information space refers to all the travel related web pages on the Internet which potential travelers can access; the interface consists of the functions of web browsers and search engines which are used to facilitate information search.

Insert Figure 1 about here

Jeng (1999) argued that travel planning follows a hierarchical structure of goals and sub-goals. It can be represented as a network of problems and sub-problems which need to be solved (Network A in Figure 2). This network of goals can be regarded as a mental model. In Network A, different nodes represent different sub-goals in which different darkness of the nodes represents different rigidity and centrality levels (the darker nodes represent more central and rigid sub-decisions). For example, "Travel Partners" is generally rigid and central to the overall travel plan, and is difficult to change; in contrast, one's choice of "Rest Stops" usually changes according to other sub-decisions of the trip. These sub-goals are interrelated and are constraints to each other. This network is the most general level of a travel information searcher's semantic mental model prior to their information search. However, this level of semantic mental model is too general and not sufficient for exploring one's semantic structure which can be compared with the language and vocabularies in the travel information space. Therefore, if one "zooms in" on the details of each node, it can be seen that each sub-goal is comprised of a cluster of related concepts, which may be nouns, adjectives, and verbs. Network B of Figure 2 is an illustration of two clusters (magnification of two nodes in Network A, Destinations and Activities). When searching for travel information on the Internet, one's choices of links are determined by the value of relevance of the link anchors; in other words, the value of information scent, which is perceived cost and value of information sources from proximal cues (Card, Pirolli, Van Der Wege et al. 2001). For example, "Theme Parks" is the most significant concept in this semantic network. If "Theme Parks", "Tropical", "South", and "Florida" appear on the same page and they are equally visually prominent, most likely the information searcher will click on the link of "Theme Parks". Additionally, not every click is equally important. Some clicks are intended to reach certain destination web pages on which the travel information searcher stays for a longer period of time. According to Kim and Hirtle (1995), information seeking on the Web involves reading/understanding and navigating, and the two processes happen simultaneously. Since attribute information regarding alternatives are needed in the decision making process, the destination page is usually a content page describing attributes of different alternatives (Nakayama, Kato, and Yamane 2000), whereas click-through web pages are index pages which serve navigational purpose. For example, a travel information searcher with a mental model as in Figure 2 will likely click through "South", "Florida", "Theme Parks" rather quickly in order to reach the "Disneyland" page, and then spend a much longer time reading its content. In this example, the former pages constitute navigational pages and the latter are reading pages. Accordingly, it is proposed that the travel information search can be represented as "episodes" whereby each episode contains one destination or content page. The rest of pages in the episode are index pages in which the travel information searcher clicks through quickly.

Insert Figure 2 about here

Importantly, during the navigation process the mental model of a travel information searcher and the representative semantic network is continuing changing. For example, after the destination choice has been made (e.g. the travel information searcher decides to go to Disneyland in Florida), some destination concepts (e.g. Disneyland and Theme Parks) in this travel planner's mind will likely transform into his/her mental peripheral and background while some related concepts will emerge and become more central (for example Hotels and Motels). Thus, travelers' mental models are dynamic and contingent upon their decision making process. After a certain period of searching and travel planning on the Internet, the information searcher will stop when all the goals are satisfied or he/she encounters obstacles (fatigue, no relevant information, or time constraint). The result of the planning effort is a "sub-space" of the entire travel information space, which represents the results of the interaction between the mental model of the travel information searcher and the travel information space. Accordingly, if the traveler's mental model and the concepts or keywords in the travel information space do not match, the information searchers will not find the information he/she seeks. Thus, mental models represent one's background knowledge, information search tasks, and one's understanding of the Internet as a travel information source and therefore, play an essential role in this conceptual framework. Interestingly, the mental model of a travel planner changes based upon the information he/she encounters and the various decisions and sub-decision he/she makes during the information search process.

METHODOLOGY

This study investigated the structure of travel information search on the Internet within the context of travel planning. In the proposed conceptual framework, the concept of mental models is a central construct in explaining both a travel information searcher's navigation behavior and his or her satisfaction with travel information search. Thus, the approach use to develop and represent travel planners' semantic models was discussed first. The second section details the thinking aloud protocol and process tracing methods which were used to capture and analyze the travel information searchers' online behavior. Last, the experimental settings of this study and detailed methods of analysis were discussed.

Representation of Travel Planners' Semantic Models

Mental models are a central concept in explaining usability when information searchers seek information in a hypertext system. However, different researchers define mental models through different perspectives. According to Norman, a mental model is "the model people have of themselves, others, the environment, and the things with which they interact. People form mental models through experience, training and instruction" (Norman 1988: 17). Furthermore, Johnson-Laird (1983) suggested that mental models are the basic structure of cognition: "… mental models play a central and unifying role in representing objects, states of affairs, sequences of events, the way the world is, and the social and psychological actions of daily life." (Johnson-Laird 1983: 397). However, mental models are incomplete and constantly evolving and usually are not accurate representations of a phenomenon (Kearsley 2001). They are parsimonious, typically contain errors and contradictions. Jacob and Shaw (1998) define a mental model as an "internal cognitive structure that the individual constructs, explicitly or implicitly, to represent a particular target domain, be it an event, an activity, an object, or a subject area." (Jacob and Shaw 1998: 158). They further argued that the concept of mental models subsumes several related constructs such as scripts (Schank and Abelson1977), schemata (Rumelhart1980), and frames (Minsky1986). Following Anderson (2000), knowledge can be divided into declarative knowledge and procedural knowledge. Declarative knowledge represents our understanding of concepts and the relationships between them; procedural knowledge stands for the knowledge of accomplishing a task. In other words, declarative knowledge is about "what" and procedural knowledge is about "how". Since the Web is mainly text-based, and Internet browsers have relatively fewer functions (bookmarks, printing, history list, and Back and Forward buttons) which are easier to learn compared with everyday uses of Internet browsers, the mental models regarding declarative knowledge is much more important in information search on the Internet. Therefore, the concept of semantic mental models is used in this research to differentiate mental models in the declarative knowledge sense (following Carley and Palmquist 1992) from traditional view of mental models in Human-Computer Interaction (HCI) (Norman 1990).

Representing Semantic Models

Connectionism is a movement in cognitive psychology which tries to explain human mind using artificial intelligent networks or similar constructs. This trend of research in cognitive psychology is in line with the macro-level of social network analysis, which treats social entities or individuals as nodes in a connected network (Wasserman and Faust1994; Doerfel and Barnett1999). As the hypertext world of the Internet changed our learning habit into non-linear and non-hierarchical, McKnight, Dillon and Richardson (1991) argue, among others, that network models are a better representation of hypertext systems and our mental constructs. The network view of cognition has been supported by various studies. The starting point of finding shared meaning and language is the understanding of cognitive process and the language comprehension (Doerfel and Barnett1999). According to Collins and Quillian (1972), human beings have a networked semantic memory, in which concepts and their relations are represented by networked nodes. Similar constructs have been proposed to explain mental models including script (Schank and Abelson1977), schema (Arbib1986), frame and k-lines (Minsky1986), which are more related to procedural knowledge. For example, Minsky proposed that the human mind is composed of networked "dummy" agents (Minsky1986), which is similar to the concept of semantic networks advanced by Collins and Quillian (1972). By combining those agents using different hierarchy and networks into a "society", Minsky demonstrated that human intelligence is possible. Even though Minsky's agent has more functionality than "concept nodes" in Quillian's semantic networks, it also demonstrated that networks are a valid representation of the knowledge state of human being's mind.

Generating Semantic Models

People have mental models which can be seen as internal representations of the world. Language is the key since languages mediate thoughts and affect categorization of concepts and even social behavior (Vygotsky1962). According to Stryker (1980), human beings respond to their symbolic world which has been classified and categorized. Accordingly, one can use language as a window to explore human beings' minds (Carley and Palmquist1992). According to Carely and Palmquist (1992), "By studying language, we can build representations of the mental models that inform social action. Moreover, through analyzing the social use of language – in both written and oral texts – we can build representations of the models that inform and shape those texts" (p. 603). Therefore, semantic network analysis on communication discourses can represent semantic mental models (Doerfel1998).

Semantic network analysis is a research method on shared meaning and a theoretical framework of analysis of cognitive process (Woelfel and Fink1980; Doerfel1998; Doerfel and Barnett1999). Research in communication states that the meaning of a concept can only be determined by the relationship with other concepts, and a model of networked concepts can accurately determine their meanings (Woelfel and Fink1980; Palmer and Barnett1984; Barnett, Palmer and Noor Al-Deen 1984; Rice and Danowski 1993). Since the concepts are embodied by the words in languages, semantic networks can represent shared semantic models and collective mental models. Semantic network analysis uses communication or discourse content to capture the relationships between different concepts (keywords) (Doerfel 1998). The content of communication message, either phone conversations (Rice and Danowki 1993), published journal articles (Doerfel and Barnett 1999), or messages in a Bulletin Board System (Danowski 1982) have been used for semantic network analysis. Traditional content analysis on texts requires human-coding (Krippendorff 1980), which involves tedious and potentially biased operations. Alternatively, semantic network analysis uses the concepts and words as they appear in the text to construct mental maps (Woelfel and Stoyanoff 1993). The advantage of this method is that there is no need for preconceived categories and the test of reliability between different human coders, and it has been used in organizational culture, communication research, marketing research and organizational structure (Doerfel and Barnett 1999).

Semantic network analysis begins with a frequency analysis on textual data which extracts the most frequently used symbols. Since the adjacency of two concepts in the text reflects their relative distance of their semantic meaning, the relative frequencies of co-occurrence of symbols represent their relative semantic distances. Neural network analysis is then used to obtain the networked mental models of symbols. Using this method, Woefel and Stoyanoff (1993) designed a program (CATPACII) to derive a semantic network from text. The tourism marketers' communication efforts, represented by the web pages they publish on the Internet, are one type of communication discourse which can be analyzed using semantic network analysis. The downloaded HTML files from these web sites can be analyzed and the generated semantic network reveals the semantic characteristics of the travel information space.

Centrality values can be calculated in order to explore different positions of keywords in a semantic network. Centrality measures how “important” or “prominent” a node is in a network (Wasserman and Faust 1994). In a nondirectional network a prominent node is the one which has many ties to other nodes, and it is more visible and involved with other concepts. Concepts with higher centrality values represent those concepts in the central positions of the semantic network. Those concepts with higher centrality value can be compared across two semantic networks. Furthermore, group centrality measures how different the concepts are in a semantic network are in terms of their diverse centrality values (Borgatti, Everett and Freeman 1999).

Measurement and Data Collection Methods

A formal travel planning exercise was designed with the following measurements and data collection methods. In the exercise, the subjects were asked to plan a trip to a designated destination using the Internet within certain period of time. To simulate a “real world” situation, a real destination was selected. The criteria for selecting the destination was: (1) this destination should have relatively diverse attractions and a variety of images; and, (2) the number of attractions should not be enormous since it will produce a huge travel information space. As such, San Diego, California in the United States was chosen as the destination for travel planning purpose.

<u>Pre-exercise Survey and Interview</u>

A pre-exercise survey was developed to collect each subject’s individual characteristics, including their travel experience, computer and Internet use experience, and the experience of using the Internet as a travel information source. The subjects’ travel experience was measured using several general questions: the number of trips in an average year; the number of trips taken in the last year; self-evaluated travel experience; and, destination-specific travel experience. Multiple questions were also used to measure the subjects' computer and Internet use experience. The objective of pre-exercise interview was to obtain data that would be used to measure the subjects' semantic mental models of San Diego prior to their travel planning. Their semantic mental models represent their search tasks, background knowledge, and their understanding of the travel information space. Following Hsieh-Yee (2001) the interview followed a semi-structured format in order to elicit diverse semantic mental models.

Online Information Search Exercise

The objective of this exercise was to examine the travel information search and travel planning process, which was defined as a goal-oriented activity and as the interaction between an Internet user and the travel information space. Prior to the exercise, the subjects were informed about the intention of this exercise and were told that they would be asked to write a short paragraph regarding their travel plan upon the completion of the exercise. The subjects were also asked to verbalize their thoughts during their information search (Ericsson and Simon 1993). One hour was set for the maximum duration of the exercise to ensure that the subjects were not exhausted and they had sufficient time to complete the travel plan. Before the formal exercise, a warm-up exercise was conducted whereby the subjects were asked to identify the weather in San Diego through the Internet. A microphone between the subject and the computer monitor was used to pick up the subject's verbalization; in addition, an online camcorder program (Camtasia) was used to record the screen activities as a movie file, including the add-on sound effects of mouse clicks and keyboard activities (TechSmith 2003). Another computer and Internet monitoring software (iOpus Starr) was used to record web pages visited, timestamps, keystrokes, and computer programs used (IOpus 2003). Last, a digital camcorder was used to record the facial expression and information behavior of the subjects; for example, sometimes the subjects wrote down information on a piece of paper or printed out web pages from the printer. Four sets of data were obtained: (1) clickstream data from the Internet monitoring software; (2) online movie recorded from the screen capturing software; (3) movies of the subjects' behavior from a digital camcorder; and, (4) artifacts from travel information search, including print-out and sheets of paper used for information organization. The combination of these four sets of data provided sufficient information for assessing the information processing during travel information search.

Semantic Network Analysis of Travel Information Space

From the initial results, it was found out that the subjects used a large amount of web sites for travel planning. The San Diego travel information space was operationally defined as those web sites which were visited by more than one subject. The web pages could be downloaded and analyzed using semantic network analysis (Woelfel and Stoyanoff 1993). The resultant semantic network, then, represents the semantic model of San Diego's travel information space.

Research Design and Procedure

Four sets of data were collected for each of the subjects including the original records from the Internet monitoring software, movies from the online screen capturing software, and digital movies from the digital camcorder, transcribed verbalization of the subjects, and information artifacts (e.g. print-out of web pages and paper used by the subjects for organizing information). The triangulation of these data sets forms an overall "protocol", including informational behavior, visited URL, subject's verbalization, and the interpreted behavior. The following analyses were conducted using this protocol.

Process-tracing and protocol analyses were first conducted to explore different information processing behavior and the structure of travel information search. Specifically, mapping the search sequence into directed graph was used to look at different episodes during users' travel information search (Hodkinson, Kiel and McColl-Kennedy 2000). In this study, click anchors were used to represent the prominent semantic concepts in an information searcher's mind. Furthermore, from information searchers' verbalization, different episodes were distinguished. For example, the information searchers may get frustrated and say, "I can't find it so I'm going to try a different site…" or "now the hotel is out of the way, let's see the parks…" Usually there is an end result to an episode: either the user got frustrated so he or she switched web sites or the user made his or her decision and he or she took on the next sub-goal. The end results were represented in boxes to distinguish them from circles which represent web nodes. Additionally, the navigation behavior was represented by labels on the paths of the graph; for example, the user may click a link to move to the next web page, or he or she can switch to a different web site by typing the URL in the browser's address box. For accesses to the same node at different times, sequential numbers were added on the path labels (see Figure 3).

Insert Figure 3 about here

In a second phase of the study, the subjects' semantic model of San Diego was compared to the semantic model of the travel information space, where it was assumed that the discrepancy and congruence between the two models reflects different views on tourism. The transcribed interviews were aggregated and analyzed using CATPACII to generate the semantic model of travel planners. The downloaded web pages from most visited tourism web sites (visited by at least two subjects) were aggregated and then analyzed to generate the semantic model of San Diego's travel information space. The number of common concepts was used as a measure of the degree of their commonality. Furthermore, QAP analysis (Quadratic Assignment Procedure) was conducted on the two matrices in order to generate the correlation (Krackhardt 1987, 1988). Thus, the two measures, number of common concepts and correlation between the two matrices, were used to measure the correlation and differences of the two semantic models.

RESULTS

The online travel exercise was conducted in August, 2002 at a large university in the United States. Fifteen subjects were recruited from various newsgroups in the university including ten undergraduate students, one graduate student, and four researchers. Their ages ranged from 19 to 45 years old with an average age of 25; six subjects were men and nine were women. Four of the subjects had been to San Diego at least once. The fifteen subjects took an average time of 36 minutes to plan the trip to San Diego, with a minimum time of 20 minutes and a maximum time of 55 minutes (the maximum time allowed was one hour). The subjects visited from 7 to 26 web sites with an average of 15. They visited from 60 to 312 web pages, averaging 124 web pages per subject. About half of the subjects (7 of 15) used a printer to print out and organize travel related information. Twelve subjects used a piece of paper to copy and organize information and two subjects used WordPad or Microsoft Word to copy and organize information.

On average, each subject took 4.2 trips in the last year and they took 2.6 trips on an average year. They considered themselves relatively experienced travelers (a mean score 5.3 on a 7-point Likert Scale question when being asked "Do you consider yourself an experienced traveler?").

Most of the subjects had been to California. Twelve subjects among the fifteen had previously used the Internet to check out destination, hotel or airline information. Fourteen of the fifteen subjects have done at least one of the following three types of activities: booking airline tickets, reserving hotel rooms or renting a car online. The subjects have used computers from 4 to 28 years with an average computer use history of 14.3 years. They used the World Wide Web from 4 to 17 years with an average of 7.5 years and they used emails from 4 o 17 years with an average of 7.7 years.

Mapping Travel Information Search on the Internet

A detailed information search protocol was generated based upon in the four sets of data described above (Table 1). Graphs representing the web search behavior were developed for each of the 15 subjects; Figure 4 provides an example of the travel information search process for Subject #13. A circle is used to represent a clicked anchor or a typed search term, indicating the dominant semantic concept in the travel information searcher's mind. At the same time, it also represents a web page resulting from the clicked link or the returned result from searched engines. A box represents the end result of one search. As can be seen, the subject found a relevant website and decided to check out one of the alternatives; he or she found some useful information, got lost, encountered a broken link, got frustrated and then decided to go back to the homepage to start the search again.

Insert Table 1 and Figure 4 about here

It is argued that the various decisions demarcate different episodes in the navigation process which can be verified from the subject's verbalization. For example, Subject #13 said: "… Now that the hotel has been taken care of… next I'm going to see what to do over there…" A summary hierarchical graph was generated from the clickstreams performed by the subjects in order to explore the detailed structure of the travel planning process (See Figure 5). The bottom level boxes contain original clickstream data whereby a sequence of clicks involving evaluation of one alternative is defined as an "episode". For Subject #13, the first two episodes focused on the Radisson Hotel and Sea World. These episodes can further be aggregated into different chapters, representing one decision and sub-problem in the subject's travel plan, e.g. hotels, attractions, or transportations. Again for Subject #13, one can see he or she moved from hotel, to attraction and shopping, and finally dinning information. However, a special episode, which is indicated by the box with a thick border in Figure 5, cannot be fit into any single type of travel plan. It is an information evaluation process in which the travel information searcher checked the address of the hotel and the attractions and located their relevant locations on the map. The travel information searcher re-evaluated the alternatives on accommodations, attractions, and dining places along the geographical frame to make his or her final decisions. The results of this analysis seems to indicate that trip planning behavior generally follows a linear fashion but with some additional re-evaluation episodes which are the exceptions of linear search.

Insert Figure 5 about here

Several major themes emerged in the generation of travel information search protocols and graphs. These findings appear to reflect the characteristics and the nature of travel information search on the Internet, not only include the structure and characteristics of online travel information search itself, but also the role of the Internet in the entire travel information environment. Information processing during travel information search on the Internet was much more diverse than anticipated. Navigation and reading processes did not represent the whole spectrum of information activities the subjects were performing online. Instead, the subjects' information behavior can be categorized into search (typing in search terms in a search box), navigation (clicking links), and information organization (printing, and writing down information on a sheet of paper) besides reading and processing attribute information for a specific alternative on web pages. Navigation can be further divided into goal-oriented navigation and browsing (Marchionini and Shneiderman 1988). Contrary to previous expectations, "reading" pages were not necessarily the pages where the subjects stayed for a longer period of time as all the pages were complex and contained large amounts of options, thereby requiring the subjects to spend time reading information and to make a decision on which link to click in order to find more detailed information. Furthermore, it was almost impossible to distinguish a navigational page from a reading page since most web pages contained a menu bar which represents the navigation structure of the web site.

More than half of subjects (8 in 15) started with hotel choices (the first chapter), and then switched to transportation or different activities and attractions. It was also found that geographic information and time information were important constraints affecting choice of hotels, means of transportation, and activities and they are not provided across different tourism web sites. The subjects always had trouble locating various hotels and attractions whereby they were forced to switch between the hotel and attraction web sites and online map web sites (e.g. MapQuest) to find geographical locations.

It appears that the subjects looked for "information portals" or "hub" pages which contain a large number of links to authoritative attractions or accommodations web pages (Kleinberg 1997).

These “information hubs” were also those web pages the subjects returned to as they started a new episode of the travel planning process. For example, the San Diego Convention and Visitors Bureau web site (http://www.sandiego.org/), the official web site for the municipal government of San Diego (http://www.sannet.gov/), and a commercial information portal for San Diego (http://www.sandiego.cc/) were chosen most often as the starting points for exploration of San Diego. This finding is consistent with information foraging theory in that information searchers tended to search information in clusters in order to minimize inter-cluster information seeking cost. Also, it confirms the validity of recent technological development of search engines which utilize information hubs and authoritative pages (Kleinberg 1997; Brin and Page 2000).

Comparing Semantic Models of Travel Information Searchers and Travel Information Space

The goal of this phase of analysis was to compute and compare semantic models of the subjects with the semantic model of the travel information space. The interviews were transcribed and aggregated across all the subjects in order to generate a semantic model of travelers. The full text of first two levels of the San Diego-related travel web sites were downloaded and analyzed using semantic network analysis (Doerfel 1998). The number of common keywords was counted and QAP analysis was conducted using the two matrices of common keywords. The travel planning exercise showed that this San Diego related travel information space is huge and highly diversified. In total, 145 unique web sites were visited by the subjects, which include not only general search engines (e.g. http://www.google.com, http://www.yahoo.com/), general tourism reservation web sites (e.g. http://www.expedia.com/, http://www.orbitz.com/, http://www.hotwire.com/), local tourism destination marketing organizations (DMO) (e.g. http://www.sandiego.org/), web sites of local attractions (e.g. http://www.cafesevilla.com/, http://www.balboapark.org/), but also information portals provided by different commercial parties (e.g. http://www.sandiego.cc/, http://www.a-zsandiego.com/) and general recreational and educational web sites (e.g. http://www.nfl.com/, http://www.nps.gov/, http://www.sdsu.edu/). Msn.com was visited by 13 of 15 subjects but it was the default page when the subjects launched the Internet Explorer; therefore, it was not counted as the most visited web site. Google was visited by 13 subjects which was the most frequently visited web site. San Diego Zoo was the most popular attraction web site which was visited by 8 subjects.

In order to obtain the semantic model of the travel information space, the first two levels of web pages of those San Diego related web sites were downloaded and analyzed using semantic network analysis. In a pilot study it was found out that downloading all the web pages on 145 web sites and analyzing them were not feasible. Instead, this research downloaded the first two levels of web pages of those web sites visited by at least two subjects (33 web sites in total). The logic for using the first two levels of web pages instead of downloading all the web pages of 33 web sites was based on three observations: (1) The first two levels of web pages are the primary information and concepts provided by tourism marketers in order to capture the attention of travelers and they represent the most prominent characteristics of the destination; (2) Downloading every web page will make the results biased toward those web sites with large quantity of web pages. However, the first pages of these 33 web sites were not necessarily directly related to San Diego. For example, the first page of yahoo.com is a general web search engine. Therefore, the first directory of San Diego was used (http://local.yahoo.com/?location_state=CA&location_city=San+Diego&location_lat=32.715710&location_lon=-117.156479). Some commercial web sites including http://www.trafficmp.com/ and http://www.mapquest.com/ were also excluded from the analysis since they have no direct connection with San Diego.

The downloaded web pages were reformatted into text files using HTML2TXT software, aggregated into one text file and analyzed using semantic network analysis (HDSE 2003). The file included 6,539 words after deleting the stop words (e.g. “and”, “or”, “is” and etc.), containing total 2,121 unique words. The top 25 keywords in the file were chosen for the analysis. In tourism information space, the top 25 keywords appeared more than 25 times. Table 2 shows the top 25 most frequently used keywords generated from semantic network analysis, including its frequencies in the file and the percentage in the total number of keywords.

Insert Table 2 about here

Figure 6 presents a two dimensional representation of the top 25 keywords generated from semantic network analysis on the travel information space. As can be seen, there are four clusters in the semantic model of the travel information space. In Quadrant I, one can see that San Diego is a “city” in “California” and the general “information” regarding “San Diego” is specifically provided on “hotel”, “park”, and “map”; in Quadrant II, the marketing focus is on water activities. San Diego has “Sea World”, “bay” area, “harbor” and “cruise”. They have the “best” “discount” “price” on “tickets”; in Quadrant III, “art” attractions are promoted; in Quadrant IV, the marketing focus of San Diego is on “shopping” “center”, “free” “event”, “museum”, “restaurant”, “center”, and “service”. These results reflect a marketing focus of online tourist information of San Diego. The four clusters reflected the popular marketing foci of travel information providers of San Diego, including water attractions, arts, general sightseeing, shopping and dining.

Insert Figure 6 about here

Centrality measures of each concept and the overall group centrality for the semantic model of the travel information space was calculated using UCINET (Borgatti, Everett and Freeman 1999; Freeman 1977). Group centrality measures the extent of difference among the centrality measures of the keywords in the semantic network where high group centrality values indicate that certain concepts have much higher degree of centrality measures than others. The concepts "Price" and "Best" had the highest centrality value, suggesting that tourism information providers' marketing focus is related to price and promotion of quality. Furthermore, popular attractions were represented in the travel information space of San Diego as represented by keywords "Sea World", "Attraction", "Center" and "Park".

Mapping of the Travel Information Searchers' Semantic Mental Model

The pre-exercise interview was semi-structured and its goal was to elicit the subject's background knowledge of San Diego and their travel preferences related to their understanding of the Internet as a travel information source. The text of the transcribed responses was aggregated together and analyzed following the same procedure described above for generating semantic model of the travel information space. The aggregated file included 1,601 keywords of which 593 were unique keywords. Again, the top 25 keywords were chosen for the analysis and comparison purpose which appeared more than 10 times (see Table 3). The top 25 keywords comprise 30.6 percent of all the keywords in the text.

Insert Table 3 about here

Similar to the semantic model of the travel information space, the subjects' semantic mental model also contained four clusters of keywords and concepts. As can be seen in Figure 7, Quadrant I includes the general travel aspects of San Diego such as "people", "beach" and "city"; other descriptive aspects of the San Diego experience include "walking around" and "live" shows. Quadrant II, on the other hand, reflects subjects' interest in various functional aspects of the trip such as "a good hotel" or a "place to stay" and "good restaurants" for "food". Interestingly, Quadrant III reflects interests in unique experiences. Last, Quadrant IV reflects the subjects' interest in "big" "attractions" like the "zoo" and "music" events. From the analysis, it is clear that the subjects have a certain level of background knowledge regarding San Diego; for example, they know that it is located in "California" and has many "beaches". They are looking for information for "hotel", "restaurant", "attractions', and they will "walk" around the "city" in the "day" and they will go to "big" "attractions", like the "zoo". Their criteria for choosing activities and attractions are "big", "good" and sometimes "different". Based upon the centrality measures of concepts in the traveler's semantic mental model, the terms "City", "Big", "Music", and "Hotel" were considered the most central concepts when planning their trip to San Diego.

Insert Figure 7 about here

Comparing Travelers' Semantic Mental Model and the Semantic model of Tourism Information Space

Direct comparison between these two semantic models can reveal differences in the languages used by tourism marketers of San Diego and by potential visitors. The common and different keywords indicate their congruence and QAP analysis on the matrices of common keywords can also reveal their relationship. A direct comparison of the two semantic models indicates that there are only 8 common concepts in the top 25 keywords (See Figure 8). The common keywords are mostly general type of concepts such as “San Diego”, “California”, “information”, “city”, “attraction”, “hotel”, “museum”, and “restaurant” and it travelers likely to locate their interested information when these keys word are used. However, the travelers’ semantic model includes more experiential and subjective concepts such as “big” and more action-based verbs such as “walk” and “look”. On the other hand, the semantic model of the travel information space is, as expected, more marketing-oriented, emphasizing pricing and products and services such as “free”, “discount”, “price”, “cruise” and “services”. The choices of attractions are also different; for example, “zoo”, “beach”, and “music” were in the top 25 concepts of the subjects’ semantic mental model whereby the travel information space includes “Sea World” as the prominent attraction. Furthermore, QAP analysis reveals no significant relationship between the two matrices on the common keywords, indicating even for the concepts in semantic models their uses and exact meaning are different.

Insert Figure 8 about here

CONCLUSIONS

The results of this study confirm that travel information search and travel planning on the Internet is a complex, dynamic and contingent process (Jeng 1999). The travel information searchers’ online behavior is much more complicated than originated expected in that the subjects performed a variety of search, navigation and information organization tasks (e.g. printing out web pages and writing on a sheet of paper), which are parts of their trip planning activity. The results of the study also indicate that trip planning on the Internet is an information-intensive task which is often beyond information searchers’ information processing capability and, therefore, is a adaptive process that mostly follows a contingent structure (Payne, Bettman and Johnson 1993; Jeng 1999). Each of the subjects appears to have a distinct semantic mental model regarding the designated destination, and each subject navigates through the travel information space according to his or her idiosyncratic semantic mental model and consequently encountered different sub-information space.

Though each subject demonstrated considerable idiosyncratic information search behavior, commonalities were also found. Eight of the fifteen subjects made accommodation choices first, indicating that the accommodation choice is a central and rigid decision of their trip plan. The results also show that information searchers generally used information hubs to facilitate the navigation process; these hubs contain many links to other authoritative pages regarding the designated destination. This result is consistent with information foraging theory (Pirolli and Card 1999) indicating that people search for information according to the significant concepts in their mental models and that people search for information hubs in order to minimize inter-cluster search cost. Information hubs could be local convention and visitor's bureau web site, commercial web sites regarding San Diego, or even a list of attractions on a local hotel web page. Furthermore, research results showed that the complex "trip-planning problem" was solved in different chapters whereby each chapter is targeting at one facet (or sub-problem) of a trip plan. Each chapter can be further divided into individual episodes in which each episode is an evaluation process of an alternative. Each subject has distinctive episodes in terms of different alternatives checked; however, on the chapter level, there are more commonalities, for example eight subjects among the fifteen made accommodation decision first.

The subjects included in this study used a large number of travel related web sites, not only including web sites provided by tourism marketers but also general web sites including media and sport web sites. The comparison between the aggregated semantic mental model of information searchers and the semantic model of the travel information space indicates that travelers have intrinsically different semantic models compared with that of travel information space. Clearly, travel information space contains a large amount of marketing-oriented information emphasizing selling including keywords such as “price”, “ticket”, and ‘free”, whereby traveler’s semantic mental models appear to be much more subjective and experiential and included the keywords “big”, “walk”, “city” and “music”. Among the top 25 most frequently used keywords, only 8 of them appeared in both semantic models, and these 8 keywords are general information keywords such as “San Diego”, “information”, “city”, “attraction” and etc. Furthermore, there are more diverse keywords in the travel information space as compared with those in the information searchers’ verbalization. These findings clearly document the different languages and lexicons the tourism information providers and the travel information searchers use.

Storytelling and narratives are a powerful way to represent information and convey meaning (Douglas 2002). Two descriptions can be generated using the top 25 words from the two semantic models by organizing those top 25 keywords into texts using as few keywords as possible and also considering the relevant distances between these keywords in Figure 6 and Figure 7. The travel information providers seem to be saying:

> **San Diego City** is located in **California**. Here we provide **information** on **hotels** and **parks**. You can also find **maps** of **San Diego**. You can find **best discount tickets** with the **best price** on **cruises**, **tours** of the **harbor**, and the **Sea World** in

> **bay** area. You can go to various **attractions** and **services**, including **shopping centers**, **museums**, **restaurants** or **free events**. We also have **best art museums**.

On the other hand, the travelers seem to perceive San Diego and the task of travel planning as:

> We know **San Diego** is a **tourist city** located in **California**. If I go there, I'd like to **walk** around on a lot of **beaches** and **look** around the **city** in the **day**. I'd like to **watch people**, and see **live** shows. I'd like to **stay** at **good hotels** and **places**, **dining** at **restaurants** with **good food**. I'd like to **look** for **information** on **good** and **different museums**. I also like to visit the **big attractions**, like the **zoo** and **music clubs**.

Clearly, the travel planners included in this study used more subjective and experiential keywords to describe their background knowledge and their information needs. In contrast, online travel information is dominated by a marketing and promotion language and focuses on profitable attractions and price information. Their points of interest are also different as shown in the descriptions. It is reasonable to assume that each party of information producers and information searchers is learning and guessing the other one's language. The information searchers would guess the keywords and concepts when he or she input the search terms into a search box to search for travel information; the tourism marketers would guess the travelers' language in order to entice them and market their destination. Thus, the results clearly demonstrated that there are still great discrepancies between these two languages. Furthermore, the keywords in the travel information searchers' semantic mental model represent their connections with designated destination. They are also the concepts and keywords associated with the destination as a brand.

Although this study provides insight into the process of travel information search on the Internet, it has many limitations related to the state of methodological development in the area under investigation. Thus, the results must be interpreted cautiously. There are at least three important limitations of this study: First, this research drew the sample from a student body in a mid-western university in the United States, and the subjects were recruited by posting advertising messages on the newsgroups. This sampling method determined that the subjects are not a representative sample of the population of all travelers in the United States. They are at a younger age, more technology-savvy, and may have a limited travel experience. Second, there are many options when conducting semantic network analysis; for example, the number of keywords to be picked in the analysis and the list of stop words used in CATPACII. The choices of these options largely depend on the goal of the research and there is no standard guidelines to follow. Even though great effort was made to ensure unbiased analysis and meaningful results, the outcomes are clearly not impartial. Last, the city (i.e., San Diego, California) was given to the participants as the designated destination for travel planning purpose. However, one might argue that the size, characteristics, and management of the city will affect its information space and therefore, the number and groups of concepts will be unique for each destination. It is argued that while exploratory in nature, this research was intended to provide insights in the travel information search on the Internet and the results provide a meaningful beginning to achieving this goal.

This research represents an initial effort in understanding the interaction between information search, travel planning and the Internet. In particular, this research explored the subjective experience of travel information search whereby no objective measures of travel information search were used. However, objective measurements of travel information search such as time spent on an information search task (efficiency) is crucial in a goal-oriented information search. Thus, future studies in this area should consider measures of efficiency which can be compared with satisfaction measure of travel information search. This research aggregated all travel web sites related to San Diego and analyzed them using semantic network analysis where the results represent the semantic model of the travel information space as a whole. Aggregated semantic model of tourism information space, however, ignore the differences of various parties comprising the tourism industry. Importantly, different parties in tourism industry might have different marketing styles and languages. Therefore, the analysis of languages of different groups of tourism parties should be conducted to determine if a better match can lead to more satisfactory travel information search experience. Last, quantitative measurement of search sequences was not considered in this study. However, exploring how different users access and navigate through travel information space using tools such as Hidden Markov Models (HMM) (MacDonald, Zucchini and Zucchii 1997) or genetic algorithms (Michell 1998) might enable one to identify the common paths of user's navigation and explore the efficiencies of travel information search. Clearly, these avenues must be considered in future research. It is hoped that this paper provides some insight into how travelers currently use the Internet, recognizing the fact that the Internet will continue to have profound impact on how travelers will plan trips in the future.

References

Allen, B. L.

1996 From Research to Design: A User-Centered Approach. *In*. Ingwersen, P., Pors, N.O. (Eds). Colis2: 2nd International Conference on Conceptions of Library and Information Science: Integration in Perspective. October 13-16, 1996: Copenhagen, Denmark (pp. 45-59). Copenhagen, Denmark: Royal School of Librarianship.

Anderson, J.R.

2000 Cognitive Psychology and Its Implications. New York: Worth Publishers.

Arbib, M.

1986 Construction of Reality. Cambridge: Cambridge University Press.

Barnett, G. A., Palmer, M. T., and Noor Al-Deen, H.

1984 An Examination of the Use of Multidimentional Scaling to Determine the Accuracy of Translation. In R. Bostrom (Eds.), Communication Yearbook, 8 (pp. 659-679). Beverly Hills: Sage.

Beaulieu, M.

2000 Interaction in Information Searching and Retrieval. Journal of Documentation 56: 431-439.

Bettman, J.R.

1979 Information Processing Theory of Consumer Choice. Boston: Addison-Wesley.

Bishop, A., and Starr, S.L.

1996 Social Informatics of Digital Library Use and Infrastructure. In M. E. Williams (Eds.) Annual review of information science and technology, 31(pp. 301-401). Medford, NJ: Information Today.

Blair, D.C.

1990 Language and Representation in Information Retrieval. Amsterdam: Elsevier Science.

Bollen, J.

2001 A Cognitive Model of Adaptive Web Design and Navigation: A shared knowledge perspective. Unpublished Dissertation, Faculty of Psychology, Free University of Brussels.

Borgatti, S. P., Everett, M.G, and Freeman, L. C.

1999 UCINET 5.0 Version 1.00. Natick: Analytic Technologies.

Brin, S. and Page, L.

2000 The Anatomy of a Large Scale Hypertextual Web Search Engine. Computer Networks and ISDN Systems 30: 107-117.

Card, S.K., Pirolli, P., Van Der Wege, M., Morrison, J.B., Reeder, R. W., Schraedley, P., and Boshart, J.

2001) Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability. In J. Jacko, A. Sears, M. Beaudouin-Lafon, and R.J.K. Jacob. (eds.) Proceedings of the ACM Conference on Human Factors in Computing Systems (pp. 498-505). ACM, Seattle, WA.

Carley, K., and Palmquist, M.

1992 Extracting, Representing, and Analyzing Mental Models. Social Forces 703: 601-636.

Chi, E.H., Pirolli, P., Chen, K., and Pitkow, J.

2001 Using Information Scent to Model User Information Needs and Actions on the Web. In Proceedings of ACM CHI 2001 Conference on Human Factors in Computing Systems (pp. 490—497). Seattle, WA: ACM Press.

Collins, A.M., and Quillian, M.R.

1972 Experiments on Semantic Memory and Language Comprehension. In L.W. Gregg (Eds.), Cognition in learning and memory (pp.117-138). New York: Wiley.

Conklin, J.

1987 Hypertext: An Introduction and Survey. Computer, III computing Society 209: 17-41.

Danowski, J. A.

1982 A Network-Based Content Analysis Methodology for Computer Mediated Communication: An illustration with a Computer Bulletin Board. In M. Burgoon (Eds.) Communication Yearbook 6 (pp. 904-925). New Brunswick: Transaction.

de Mey, M.

1977 The Cognitive View Point: Its Development and Its Scope. In: de Mey, M., Pinxten, R, Poriau, M., Vandamme, F. (Eds) CC77: International Workshop on the Cognitive ViewPoint, March 24-26, 1977: Ghent, Belgium (pp. 7-23). Ghent, Belgium: University of Ghent: xvi-xxxi.

Doerfel, M.

1998 What Constitutes Semantic Network Analysis? A Comparison of Research and Methodologies. Connections 212: 16-26.

Doerfel, M.L. and Barnett, G.A.

1999 A Semantic Network Analysis of the International Communication Association. Human Communication Research 254: 589-603.

Douglas, R.A.

2002 How Storytelling Builds Next-Generation Leaders. MIT Sloan Management Review 43: 63-69.

Engel, J.F., Blackwell, R.D., and Miniard, P.W.

1990 Consumer Behavior. Orlando, FL: The Dryden Press.

Ericsson, K. A., and Simon, H. A.

1993 Protocol Analysis: Verbal Reports as Data. Cambridge, MA: The MIT Press.

Fodness, D., and Murray, B.

1999 A Model of Tourist Information Search Behavior. Journal of Travel Research 37: 220-230.

Freeman, L.C.

1977 A Set of Measures of Centrality Based on Betweenness. Sociometry 40: 35-41.

Furnas, G.W., Landauer, T.K., Gomez, L.M., and Dumais, S.T.

1987 The Vocabulary Problem in Human-System Communication. Communications of the ACM 30: 964—971.

HDSE.

2003 HTML2TXT Documentation. Available Online at: http://hdse.org/html2txt/documentation/GeneralHelp.htm.

Hodkinson, C., Kiel, G., and McColl-Kennedy, J. R.

2000 Consumer Web Search Behaviour: Diagrammatic Illustration of Wayfinding on the Web. International Journal of Human Computer Studies 52: 805-830.

Hsieh-Yee, I.

2001 Research on Web Search Behavior. Library and Information Science Research. 23: 167-185.

Hwang, Y.H., and Fesenmaier, D.R.

2001 Collaborative Filtering: Strategies for Travel Destination Bundling. Sheldon, Wöber and Fesenmaier (Eds.), Proceedings of ENTER' 2001 in Montreal, Canada (pp. 358 – 267). New York: Springer.

Ingwersen, P.

1992 Information Retrieval Interaction. London: Taylor Graham.

IOpus.

2003 STARR PC & Internet Monitor User Manual. IOpus. Available Online at: http://www.iopus.com/starr/index.htm.

Jacob, E.K., and Shaw, D.

1998 Sociocognitive Perspectives on Representation. Annual Review of Information Science and Technology 33: 131-185.

Jeng, J.

1999 Exploring the Travel Planning Hierarchy: An Interactive Web Experiment. Unpublished Dissertation. University of Illinois at Urbana-Champaign.

Johnson-Laird, P.N.

1983 Mental models: Toward a Cognitive Science of Language, Inference, and Consciousness. Cambridge: Harvard University Press.

Kearsley, G.

2001 Mental Models. In Encyclopedia of Psychology, retrieved July 16, 2001 from http://tip.psychology.org/models.html.

Kim, H., and Hirtle, S.C.

1995 Spatial Metaphors and Disorientation in Hypertext Browsing. Behavior and Information Technology 45: 93-102.

Kleinberg, J.

1997 Authoritative Sources in a Hyperlinked Environment. Proceeding of the 9th ACM-SIAM Symposium on Discrete Algorithms.

Klicek, B.

2001 Tourist's Decision Making Process Assisted by the Web and Multimedia Intelligent Advisory System. Sheldon, Wöber and Fesenmaier (Eds.), Proceedings of ENTER' 2001 in Montreal, Canada (pp. 358 – 267). New York: Springer.

Krackhardt, D.

1987 QAP Partialling as a Test Spuriousness. Social Networks 9: 171-186.

Krackhardt, D.

1988 Predicting with Networks: A Multiple Regression Approach to Analyzing Dyadic Data. Social Networks 10: 359-381.

Krippendorff, K.

1980 Content Analysis: An Introduction to Its Methodology. Newbury Park: Sage Publications.

Lake, D.

2001 American Go Online for Travel Information. CNN, June 14, 2001. Available: http://www.cnn.com/2001/TECH/internet/06/14/travelers.use.net.idg/index.html

Linden, G., Hanks, S., and Lesh, N.

1997 Interactive Assessment of User Preference Models: The Automated Travel Assistant. in Proceedings of the 6th International Conference on User Modeling (UM97) , June 2-5, CISM, Vol. 383 (pp. 67-78). New York: Springer.

MacDonald, I. L., Zucchini, W., Zucchii, W.

1997 Hidden Markov and Other Models for Discrete- Valued Time Series. Boca Raton, FL: CRC Press.

McKnight, C., Dillon, A., and Richardson,J.

1991 Hypertext in Context. Cambridge: Cambridge University Press.

Michell, M.

1998 An Introduction to Genetic Algorithms (Complex Adaptive Systems), Cambridge: MIT Press.

Minsky, M.

1986 The society of Mind. New York: Simon and Schuster.

Morkes, J., and Nielsen, J.

1997 Concise, SCANNABLE, and Objective: How to Write for the Web. Available online: http://www.useit.com/papers/webwriting/writing.html.

Nakayama, T., Kato, H., and Yamane, Y.

2000). Discovering the Gap between Web Site Designers' Expectations and Users' Behavior, Computer Networks 33: 811-822.

Newhagen, J.E., and Rafaeli, S.

1996 Why Communication Researchers Should Study the Internet: A Dialogue. Journal of Computer-Mediated Communication[On-line], 1. Available: http://www.ascusc.org/jcmc/vol1/issue4/rafaeli.html.

Newman, J.W.

1977 Consumer External Search: Amounts and Determinants. In A. Woodside, J. Sheth and P. Bennett (Eds.) Consumer and Industrial Buying Behavior (pp. 79-94). New York: North-Holland.

Norman, D.A.

1988 The Psychology of Everyday Things. New York: Basic Books.

Norman, D.A.

1990 The Design of Everyday Things. New York: Doubleday Publishing.

Palmer, M.T., and Barnett, G.A.

1984 Using Spatial Models to Determine the Accuracy of Language Translation. In W.B. Gudykunst and Y. Y. Kim (Eds.), International and Intercutural Communication Annual, 18 (129-147). Thousand Oaks: Sage.

Pan, B., and Fesenmaier, D.R.

2000 A Typology of Tourism-Related Web Sites: Its Theoretical Background and Implications. Information Technology and Tourism 3: 155-176.

Payne,J.W., Bettman J.R., and Johnson, E.J.

1993 The Adaptive Decision Maker. Cambridge: University of Cambridge Press.

Pirolli, P., and Card, S.K.

1999 Information Foraging. Psychological Review 106: 643-675.

Poon, A.

1993 Tourism, Technology and Competitive Strategies. Oxon: CAB International.

Radosevich, L.

1997 Fixing Web-site Usability. InfoWorld 19: 81-82.

Ratchford, B.T., Talukdar, D., and Lee, M. S.

2001 A Model of Consumer Choice of the Internet as an Information Source. International Journal of Electronic Commerce 5: 7-21.

Ricci, F., and Werthner, H.

2001 Case-based Destination Recommendations over an XML Data Repository. Sheldon, Wöber and Fesenmaier (Eds.), Proceedings of ENTER' 2001 in Montreal, Canada (pp. 150 – 159). New York: Springer.

Rice, R.E. and Danowski, J.

1993 Is It Really Just Like a Fancy Answering Machine? Comparing semantic networks of different types of voice mail users. Journal of Business Communication 30: 369-397.

Rich, C., and Sidner, C.L.

1998 COLLAGEN: A Collaboration Manager for Software Interface Agents. Working paper, http://www.merl.com/reports/TR97-21a/index.html

Rumelhart, D.E.

1980 Schemata: The Building Blocks of Cognition. In: Spiro, R. J., Bruce, B.C., Brewer, W.F. (Eds). Theoretical Issues in Reading Comprehension: Volume 1 (pp. 38-58). Hillsdale: Lawrence Erlbaum.

Schank, R., and Abelson, R.P.

1977 Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures. Hillsdale: Lawrence Erlbaum.

Shera, J.H.

1965 Putting Knowledge to Work. In Shera, J.H. (Eds). Libraries and Organization of Knowledge (pp. 129-142). Hamden, CT: Archon Books.

Stoltz, C.

1999 Each Year, a Bit Less. Washington Post (1999, Nov 11).

Stryker, S.

1980 Symbolic Interactionism. Menlo Park, CA: Benjamin/Cummings.

TechSmith.

2003 Camtasia Studio Documentation. TechSmith, Available Online at: http://wcarchive.cdrom.com/pub/bws/bws_48/studioug.pdf.

Vogt, C.A., and Fesenmaier, D.R.

1998 Expanding the Functional Information Search Model. Annals of Tourism Research 25(3), 551-578.

Vygotsky, L.

1962 Thought and Language. New York: Wiley.

Wasserman, S., and Faust, K.

1994 Social Network Analysis. Cambridge: Cambridge University Press.

Woelfel, J., and Fink, E.L.

1980 The Measurement of Communication Processes: GALILEO Theory and Method. New York: Academic Press.

Woelfel, J., and Stoyanoff, N. J.

1993 CATPAC: A Neural Network for Qualitative Analysis of Text. Paper presented at the annual meeting of the Australian Marketing Association, Melbourne, Australia.

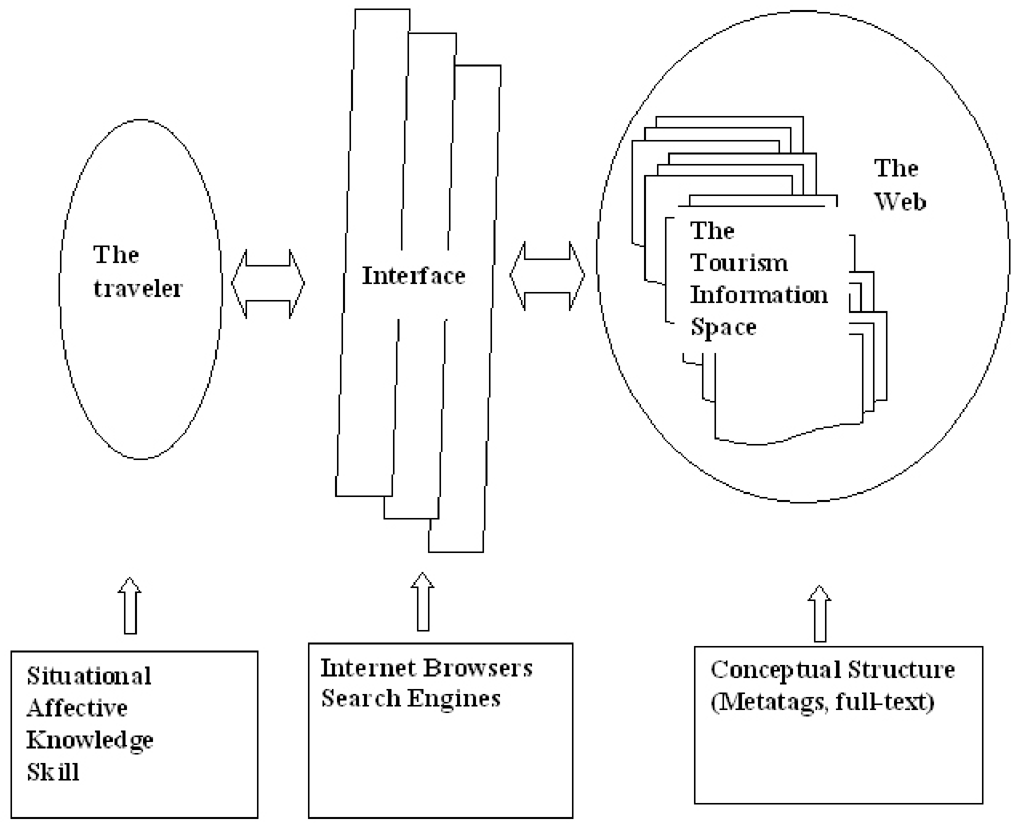

Figure 1. Travelers Interacting with the Internet When Searching Information

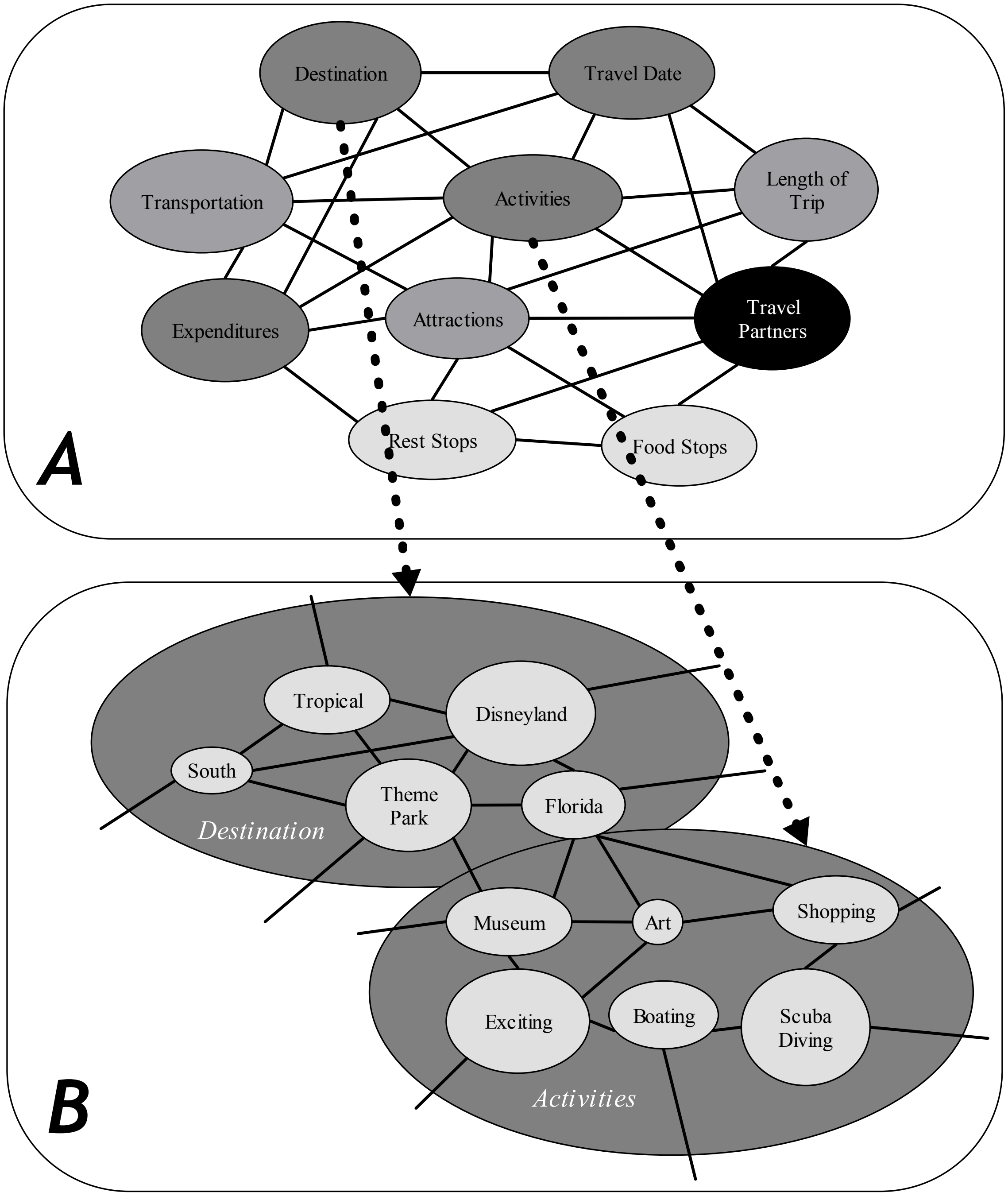

Figure 2. An Illustrative Semantic mental Model of Travel Planning

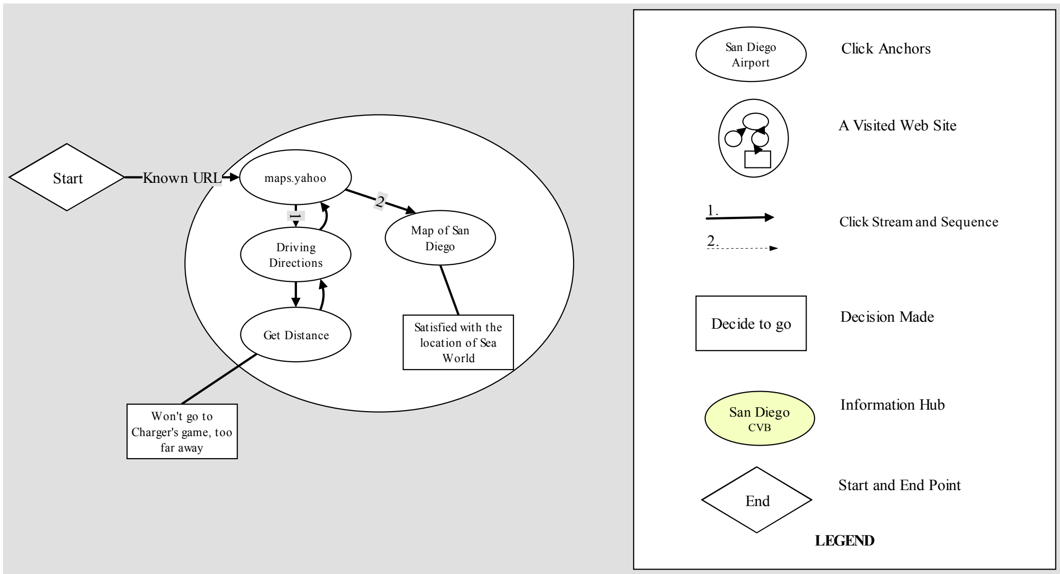

Figure 3. Example Graph of the Information Search Process

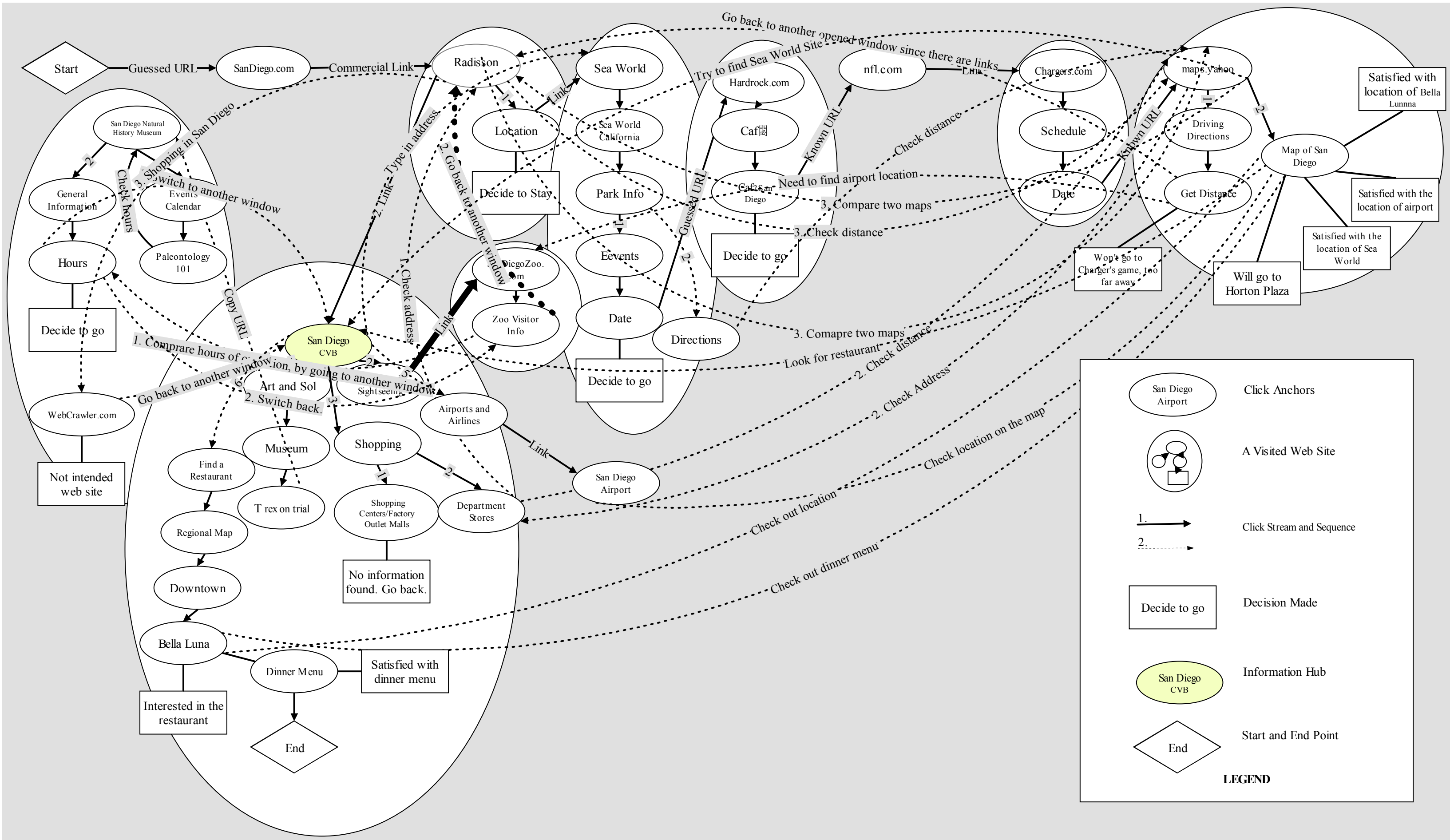

Figure 4. Clickstream Semantic Map of Subject #13

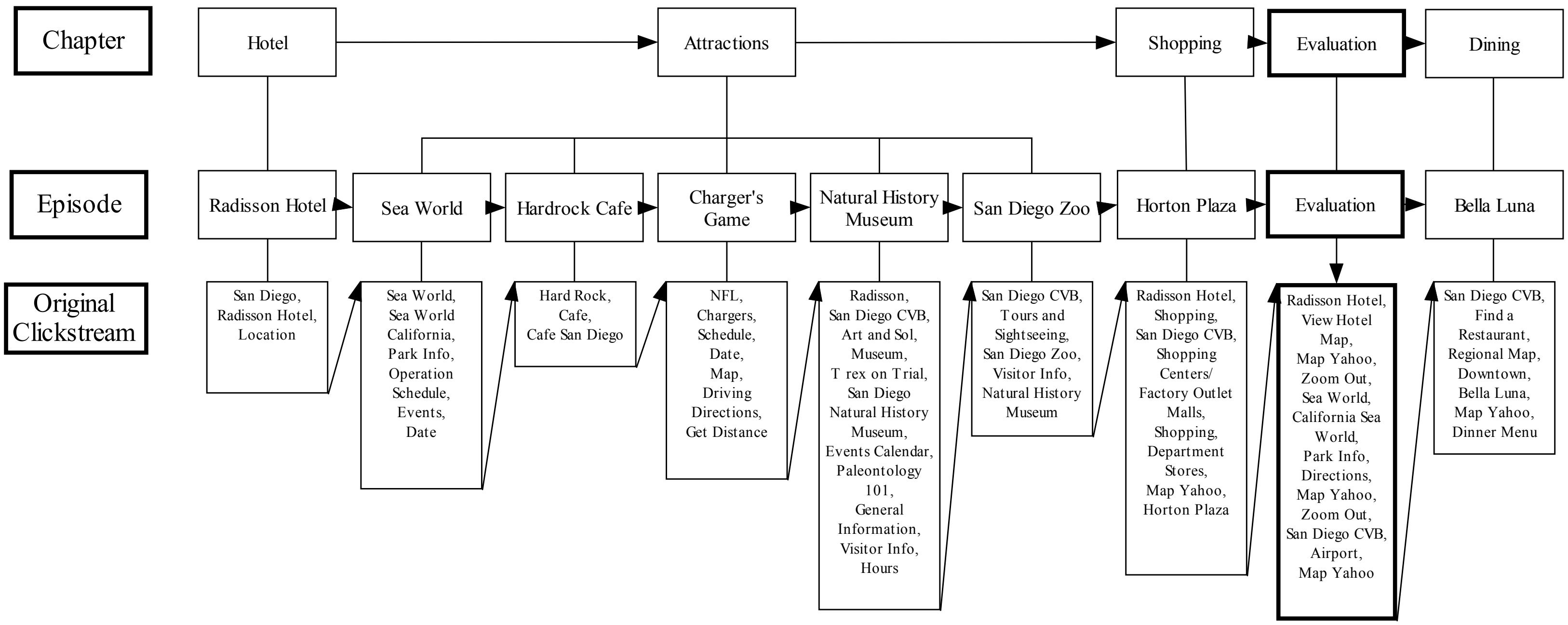

Figure 5. Deconstruction of Clickstream Semantic Map of Subject #13

Figure 6. Semantic Mapping of San Diego Tourism Information Space

Figure 7. Semantic Mapping of Travel Planners to San Diego

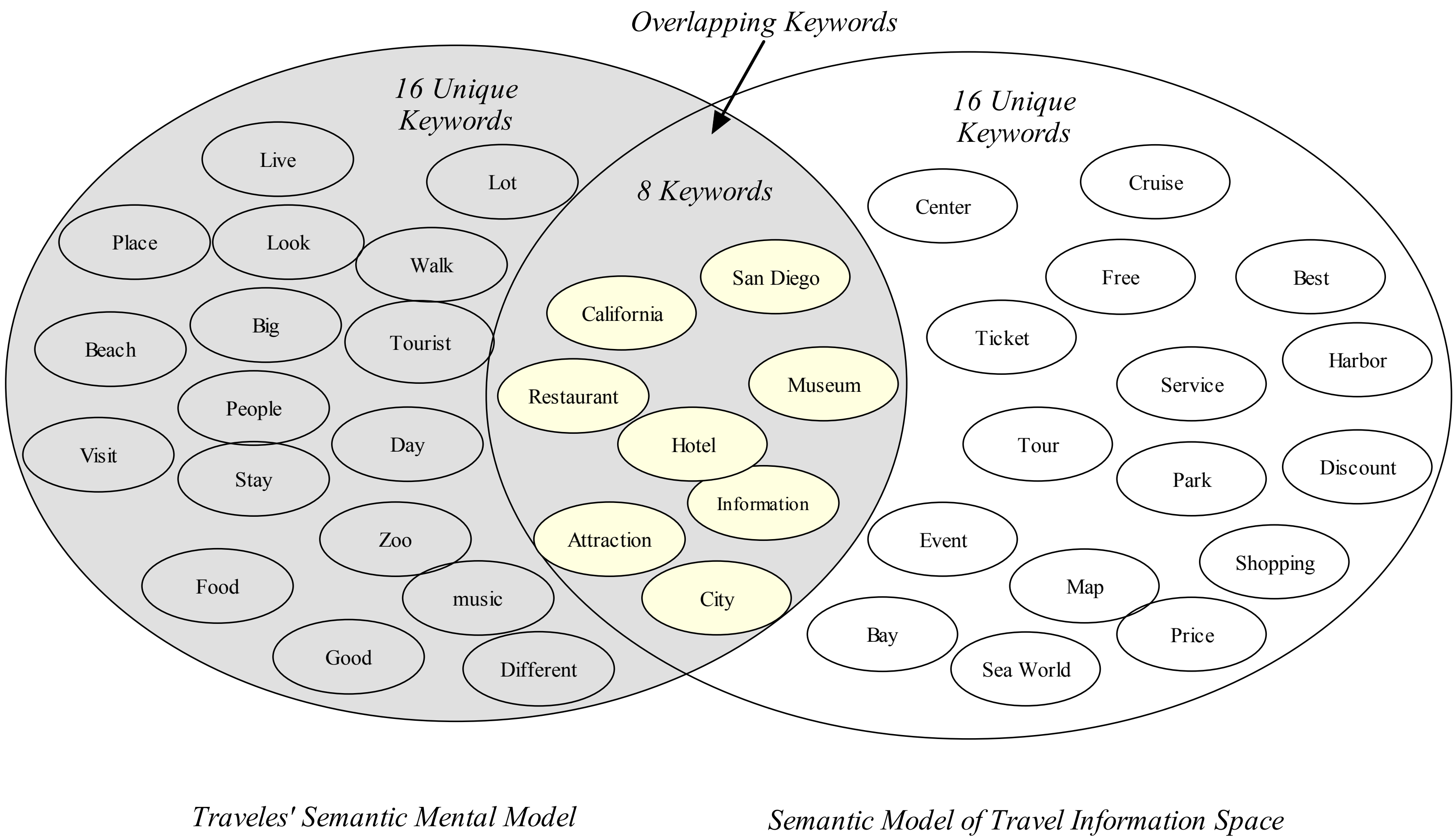

Figure 8. Comparison of Two Semantic Models of San Diego

Time Spent	Time	Behavior	Verbalization	Interpretation
0:00:03	20:15:42	Click Link	Go back to use Internet Explorer,	Start IE.
0:00:09	20:15:45	Type In	I'll go to google… SanDiego.com to see what I find here. See if there's anything interesting from the web site I can find out. Since I haven't been to the city.	Type in guess address: **sandiego.com**
0:00:20	20:15:54	Click Link	All right, so we're here. Just looking around to see what's on here. Hotels, all right. Radisson Hotel San Diego. That catches my eye.	Click on **Radisson Hotel San Diego**.
0:00:11	20:16:14	Click Link	Stayed at Radisson before. I like them. Looks pretty nice. 89 to 90. I can deal with that for a day.	Scan information and click on **location**.
0:00:31	20:16:25	Click Link	All right, location. Says, the heart of San Diego. Minutes from the major attractions. That's a good thing. Oh, Sea World. Oh, I probably make time to do that. I want to go to Sea World. I like fish. Working at marine labs, so. That'll be something I'm very interested in doing. All right. It's minutes, it says, to the attractions.	Click on **Sea World** link.
0:00:06	20:16:56	Click Link	I probably just take a taxi and not worry about renting a car for a day. It's a little bit too much.	Click on **California** on Sea World page.
0:00:10	20:17:02	Click Link	Too much work. All right. I'm at Sea World. So let's see. Park info.	Click on Park Info.
0:00:12	20:17:12	Click Link	Let's see. Wow, they have varieties, too. Hours of operation. Wow. It's expensive.	Check out detailed information and click on **Operation Schedule.**
0:00:03	20:17:24	Click Link	…	Choose month.
0:00:22	20:17:27	Click Link	Choose the month of October. I'm gonna be there 19th and 20th. So it's open until 7:30PM, open at 10AM. Sounds like a good way to spend my Saturday once I get there. Check in my hotel and go to the Sea World. Really great idea. That's what I wanna do Saturday.	Look at the **schedule** and decide to go there on Saturday.
0:00:04	20:17:49	Type In	Let me check hard rock café dot com.	Type in address of **Hard Rock café**.

Table 1. Example of Information Search Protocol

Table 2. Tourism Information Space for San Diego

Web Site	Number of Subjects Visited	Web Site	Number of Subjects Visited
www.google.com	13	www.sannet.gov	2
www.sandiegozoo.org	8	www.sandiego-online.com	2
www.sandiego.org	7	www.sandiego.cc	2
www.sdcommute.com	5	www.revup.biz	2
www.mapquest.com	5	www.reservetravel.com	2
www.trafficmp.com	4	www.portofsandiego.org	2
www.seaworld.com	3	www.orbitz.com	2
www.sandiego.com	3	www.netster.com	2
www.expedia.com	3	www.infosandiego.com	2
www.blueescape.com	3	www.hotwire.com	2
www.a-zsandiegoattractions.com	3	www.fodors.com	2
www.yahoo.com	2	www.citysearch.com	2
www.thebigbay.com	2	www.cafesevilla.com	2
www.sdsu.edu	2	www.balboapark.org	2
www.sdro.com	2	www.arestravel.com	2
www.sdnhm.org	2	www.4adventure.com	2
www.sdinsider.com	2		

Table 3. Top 25 Keywords in the Semantic model of Online Tourism Information Space

Words	Frequency	Percentage	Words	Frequency	Percentage
San Diego	241	8.2	Free	31	1.1
Ticket	80	2.7	Best	30	1.0
Information	62	2.1	Attraction	29	1.0
Hotel	59	2.0	Cruise	29	1.0
Center	57	1.9	Art	28	1.0
Tour	53	1.8	Discount	28	1.0
Event	52	1.8	Map	28	1.0
City	51	1.7	Museum	28	1.0
Park	43	1.5	Sea World	28	1.0
Restaurant	43	1.5	Harbor	27	0.9
Service	39	1.3	Bay	26	0.9
California	38	1.3	Shopping	26	0.9
Price	37	1.3			

Table 4. Top 25 Keywords in Travelers' Semantic Mental Model

Words	Frequency	Percentage	Words	Frequency	Percentage
San Diego	40	3.7	Stay	16	1.5
Place	37	3.4	Walk	16	1.5
Beach	35	3.2	Look	14	1.3
Museum	30	2.8	Tourist	14	1.3
Food	27	2.5	Zoo	14	1.3
California	25	2.3	Day	13	1.2
Restaurant	23	2.1	Big	12	1.1
Hotel	22	2.0	Different	12	1.1
Information	21	1.9	Good	12	1.1
Lot	21	1.9	Visit	12	1.1
City	19	1.8	Attraction	11	1.0
People	17	1.6	Music	11	1.0
Live	16	1.5			