Travel Information Search on the Internet: An Exploratory Study

Bing Pan
National Laboratory for Tourism and eCommerce
Department of Leisure Studies
University of Illinois at Urbana-Champaign
Introduction

**Overall Structure**
- Background/Purpose of Study
- Assumptions
- Conceptual Framework
- **Methodology**
- Results
- Conclusions/Implications
Background

- 95 percent of web users used the Internet to gather travel related information; 93 percent use the Web for planning vacations (Lake, 2001)
- Frustrating travel planning on the Internet (Stoltz, 1999)
- Complex, contingent, and dynamic travel planning (Jeng, 1999)
- The vocabulary problem (Furnas, Landauer, Gomez & Dumais, 1987)
- We don’t know how users used the Internet for travel planning
- Research on travel information search on the Internet is scarce; most are survey research
Purpose of Research

- Explore the structure of travel information search on the Internet in the context of travel planning
- Examine the satisfaction of travel information search on the Internet and its determinants
Travel Planning and Information Search

- Travel planning is (Jeng, 1999):
  - A hierarchy of decisions which involves a set of sub-decisions, for example, destination, travel partners, accommodation, dining and others;
  - Multi-facet, dynamic and contingent process.
Core Assumptions 2

Usability of Information System

- The mismatch between user’s mental model and system’s conceptual model contributes to usability problem (Norman, 1986);

- The polyrepresentation of concepts in the languages of the users’ cognitive space and the information system is a major issue when designing an effective information system interface (Ingwersen, 1996).

- Mismatch between travel information searcher’s mental model and the semantic model of travel information space.
Core Assumptions 3

Mental Models and Information Search

- Mental model as intermediate construct
- Semantic mental model vs. mental model in HCI
- Declarative Knowledge vs. Procedural Knowledge (Anderson, 2000)
- Using semantic networks to represent semantic mental models (Collins and Quillian, 1972; Doerfel, 1998)
Traveler's Semantic Mental Model
A Conceptual Model

- Choices of links are based on the relative value of information scent: the link anchors (text or pictures)
- Navigation process and reading process
- Search process can be broken up into different episodes; each episode targets at a sub-problem
Research Methodology

Travel Planning Exercise on the Internet
Research Methods

1. **Thinking-aloud Protocol:** Capturing information processing during the process

2. **Semantic network analysis and QAP analysis:**
   Using transcripts from interviews and text from full texts of web pages
   1. Transcripts from interviews
   2. Full texts of web pages from visited web sites

3. **Correlation analysis:** Using survey data and results from semantic analysis
Research Methodology

Phase I: Explore Structure

- Clickstream
- Screen Capturing
- Artifacts
- Behavior Movie

Data Triangulation

Information Search Protocol

Concept Mapping

Navigation Graph

Translation

Hierarchical Graph
Phase II. Global Level of Comparisons

**Research Methodology**

- **Travelers' Semantic Mental Model**
  - Pre-Exercise Interview
  - Aggregation
  - Semantic Network Analysis
    - Number of Common Keywords & QAP Analysis on Common Keywords

- **Travel Information Space**
  - Full Text of Web Sites
  - Semantic Network Analysis

**National Laboratory for Tourism and eCommerce**
**University of Illinois at Urbana-Champaign**
Research Methodology

Phase III: Examine Satisfaction

Overall Satisfaction Model
Analysis and Results
## Phase I: Information Search Protocol: Data

<table>
<thead>
<tr>
<th><strong>Clickstream</strong></th>
<th><strong>Online Activity</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(Internet Monitoring Software)</td>
<td>(Screen Capturing Software)</td>
</tr>
<tr>
<td><strong>Artifacts from Information Search</strong></td>
<td><strong>Large Scale Behavior</strong></td>
</tr>
<tr>
<td></td>
<td>(Digital Camcorder)</td>
</tr>
</tbody>
</table>
## Phase I: Original Clickstream

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Duration</th>
<th>Type</th>
<th>URL and Details</th>
<th>Folder Path</th>
<th>Target URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/06/2002</td>
<td>20:15:40</td>
<td>00:01</td>
<td>TITLE</td>
<td>Program Manager</td>
<td><code>c:\winnt\explorer.exe</code></td>
<td>-</td>
</tr>
<tr>
<td>08/06/2002</td>
<td>20:15:41</td>
<td>00:04</td>
<td>TITLE</td>
<td>Welcome to MSN.com - Microsoft Internet Explorer</td>
<td><code>c:\program files\internet explorer\iexplore.exe</code></td>
<td>-</td>
</tr>
<tr>
<td>08/06/2002</td>
<td>20:15:42</td>
<td></td>
<td>Link</td>
<td>Welcome to MSN.com - Microsoft Internet Explorer</td>
<td><code>c:\program files\internet explorer\iexplore.exe</code></td>
<td><code>http://www.msn.com/</code></td>
</tr>
<tr>
<td>08/06/2002</td>
<td>20:15:45</td>
<td></td>
<td>Keystrokes</td>
<td>Welcome to MSN.com - Microsoft Internet Explorer</td>
<td><code>c:\program files\internet explorer\iexplore.exe</code></td>
<td>w</td>
</tr>
<tr>
<td>08/06/2002</td>
<td>20:15:45</td>
<td>02:04</td>
<td>SUBTITLE</td>
<td>Welcome to MSN.com - Microsoft Internet Explorer</td>
<td><code>c:\program files\internet explorer\iexplore.exe</code></td>
<td>-</td>
</tr>
<tr>
<td>08/06/2002</td>
<td>20:15:45</td>
<td></td>
<td>Keystrokes</td>
<td>Welcome to MSN.com - Microsoft Internet Explorer</td>
<td><code>c:\program files\internet explorer\iexplore.exe</code></td>
<td><code>http://www.sandiego.com</code></td>
</tr>
</tbody>
</table>
## Phase I: Information Search Protocol: Data

<table>
<thead>
<tr>
<th>Clickstream</th>
<th>Online Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Internet Monitoring Software)</td>
<td>(Screen Capturing Software)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Artifacts from Information Search</th>
<th>Large Scale Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Digital Camcorder)</td>
<td>(Digital Camcorder)</td>
</tr>
</tbody>
</table>
Phase I: Information Search Artifacts

Comfort Inn hotel circle
2201 HOTEL CIRCLE SOUTH
SAN DIEGO, CA 92108
HOTEL CIRCLE

art museum
1450 El Prado,
Balboa Park,
San Diego, California

Coronado
1100 Orange Ave,
Coronado.

wild animal park
San Pasqual Valley Road

Escondido CA 92027-7017
### Phase I: Information Search Protocol: Data

<table>
<thead>
<tr>
<th>Clickstream</th>
<th>Online Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Internet Monitoring Software)</td>
<td>(Screen Capturing Software)</td>
</tr>
<tr>
<td>Artifacts from Information Search</td>
<td>Large Scale Behavior</td>
</tr>
<tr>
<td></td>
<td>(Digital Camcorder)</td>
</tr>
</tbody>
</table>
### Phase I: Final Protocol

<table>
<thead>
<tr>
<th>Time Spent</th>
<th>Time</th>
<th>Behavior</th>
<th>Verbalization</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:00:03</td>
<td>20:15:42</td>
<td>Click Link</td>
<td>Go back to use Internet Explorer,                                                                ----------------------------------------------------------------------------------------------------------</td>
<td>Start IE.</td>
</tr>
<tr>
<td>0:00:09</td>
<td>20:15:45</td>
<td>Type In</td>
<td>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.</td>
<td>Type in guessed address: sandiego.com</td>
</tr>
<tr>
<td>0:00:20</td>
<td>20:15:54</td>
<td>Click Link</td>
<td>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.</td>
<td>Click on Radisson Hotel San Diego.</td>
</tr>
<tr>
<td>0:00:11</td>
<td>20:16:14</td>
<td>Click Link</td>
<td>Stayed at Radisson before. I like them. Looks pretty nice. 89 to 90. I can deal with that for a day.</td>
<td>Scan information and click on Location.</td>
</tr>
<tr>
<td>0:00:31</td>
<td>20:16:25</td>
<td>Click Link</td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td>0:00:06</td>
<td>20:16:56</td>
<td>Click Link</td>
<td>I probably just take a taxi and not worry about renting a car for a day. It's a little bit too much.</td>
<td>Click on California on Sea World page.</td>
</tr>
<tr>
<td>0:00:10</td>
<td>20:17:02</td>
<td>Click Link</td>
<td>Too much work. All right. I'm at Sea World. So let's see. Park info.</td>
<td>Click on Park Info.</td>
</tr>
<tr>
<td>0:00:12</td>
<td>20:17:12</td>
<td>Click Link</td>
<td>Let's see. Wow, they have varieties, too. Hours of operation. Wow. It's expensive.</td>
<td>Check out detailed information and click on Operation Schedule.</td>
</tr>
</tbody>
</table>
Phase I: A Navigation Map
Phase I: A Hierarchy of Travel Planning Process

National Laboratory for Tourism and eCommerce
University of Illinois at Urbana-Champaign
Phase I: Findings

– Hierarchical structure

– Information behavior: searching, navigation and information organization

– Information hubs

– Accommodation: the most central sub-decision
Phase I: Other Findings

- Combination of information sources and decision aids
- Geographical information and time frame
- Non-searchable attributes
## Phase II. Travel Information Space

<table>
<thead>
<tr>
<th>Web Site</th>
<th>Number of Subjects Visited</th>
<th>Web Site</th>
<th>Number of Subjects Visited</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.google.com">www.google.com</a></td>
<td>13</td>
<td><a href="http://www.sannet.gov">www.sannet.gov</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sandiegozoo.org">www.sandiegozoo.org</a></td>
<td>8</td>
<td><a href="http://www.sandiego-online.com">www.sandiego-online.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sandiego.org">www.sandiego.org</a></td>
<td>7</td>
<td><a href="http://www.sandiego.cc">www.sandiego.cc</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sdcommute.com">www.sdcommute.com</a></td>
<td>5</td>
<td><a href="http://www.revup.biz">www.revup.biz</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.mapquest.com">www.mapquest.com</a></td>
<td>5</td>
<td><a href="http://www.reservetravel.com">www.reservetravel.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.trafficmp.com">www.trafficmp.com</a></td>
<td>4</td>
<td><a href="http://www.portofsandiego.org">www.portofsandiego.org</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.seaworld.com">www.seaworld.com</a></td>
<td>3</td>
<td><a href="http://www.orbitz.com">www.orbitz.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sandiego.com">www.sandiego.com</a></td>
<td>3</td>
<td><a href="http://www.netster.com">www.netster.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.expedia.com">www.expedia.com</a></td>
<td>3</td>
<td><a href="http://www.infosandiego.com">www.infosandiego.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.bluescape.com">www.bluescape.com</a></td>
<td>3</td>
<td><a href="http://www.hotwire.com">www.hotwire.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.a-zsandiegoattractions.com">www.a-zsandiegoattractions.com</a></td>
<td>3</td>
<td><a href="http://www.fodors.com">www.fodors.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.yahoo.com">www.yahoo.com</a></td>
<td>2</td>
<td><a href="http://www.citysearch.com">www.citysearch.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.thebigbay.com">www.thebigbay.com</a></td>
<td>2</td>
<td><a href="http://www.cafesevilla.com">www.cafesevilla.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sdsu.edu">www.sdsu.edu</a></td>
<td>2</td>
<td><a href="http://www.balboapark.org">www.balboapark.org</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sdro.com">www.sdro.com</a></td>
<td>2</td>
<td><a href="http://www.arestravel.com">www.arestravel.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sdnhm.org">www.sdnhm.org</a></td>
<td>2</td>
<td><a href="http://www.4adventure.com">www.4adventure.com</a></td>
<td>2</td>
</tr>
<tr>
<td><a href="http://www.sdinsider.com">www.sdinsider.com</a></td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Phase II. Comparison of Two Semantic Models

Travelers' Semantic Mental Model

Semantic Model of Travel Information Space

Overlapping Keywords

16 Unique Keywords

8 Keywords

Look

Big

music

Hotel

City

Discount

Service

Park

Discount

Best

Cruise

Center

Ticket

Tour

Park

Shopping

Sea World

Event

Map

Bay

Price

Good

Different

Zoo

Attraction

Information

San Diego

Restaurant

Museum

California

Tourist

Day

Walk

Place

Live

Lot

16 Unique Keywords

Visit

Stay

People

Beach

Look

Big
Phase II. Semantic Model of Travel Information Space
Phase II. Traveler’s Semantic Mental Model
Phase II: Findings

- Huge and diverse travel information space

- Different languages
  - Marketing language, price and quality
  - Traveler’s language, subjective and experiential

- Common concepts are general concepts about information, hotel, restaurant, etc.
San Diego City is located in California. Here we provide information on hotels and parks. You can also find maps of San Diego. Here you can find best discount tickets with the best price on cruises, tours of 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.

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.
### Phase III: Congruence of Individual Model with Information Space

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Number of Common Concepts</th>
<th>Ratio of Common Concepts</th>
<th>QAP Correlation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>0.20</td>
<td>0.39</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.20</td>
<td>0.80</td>
<td>0.33</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>0.17</td>
<td>0.28</td>
<td>0.12</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>0.41</td>
<td>0.28</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0.10</td>
<td>0.30</td>
<td>0.26</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0.33</td>
<td>0.48</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>0.16</td>
<td>0.31</td>
<td>0.03</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>0.17</td>
<td>0.10</td>
<td>0.50</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>0.24</td>
<td>0.36</td>
<td>0.11</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0.14</td>
<td>0.33</td>
<td>0.03</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>0.26</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>0.11</td>
<td>0.24</td>
<td>0.16</td>
</tr>
<tr>
<td>13</td>
<td>8</td>
<td>0.24</td>
<td>0.35</td>
<td>0.01</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>0.18</td>
<td>0.44</td>
<td>0.05</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>0.09</td>
<td>0.13</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>6.3</strong></td>
<td><strong>0.20</strong></td>
<td><strong>0.33</strong></td>
<td>--</td>
</tr>
</tbody>
</table>
Phase III. Examination of Satisfaction Model

Travel Experience

Computer and Internet Use Experience

Internet Use Experience as Travel Information Source

Travel Information Searcher's Model

Tourism Information Space's Model

Outcome Satisfaction

Negatively Correlated

Positively Correlated
Phase III. Examination of Satisfaction Model

(travel planning) “...is easy. Everything is settled, set. I have an itinerary set up already for me. I know when I am leaving, where my car will be at, which hotel I will be staying at, everything is reserved. I probably call a couple days before. I probably call the hotel and probably call the car rental company just to double-check, but basically everything is set up. Everything is paid for. Why I am satisfied? Not worried.”

“Well, I was satisfied, but not completely satisfied, because my ideal planning time for a vacation will probably be like a couple of hours. Maybe between two to five hours’ research. Maybe looking at as much details as I can before making a decision. I feel probably this was as good as I could have done in the time set. I would like to have more time.”

“...Well, surprises all are found out. I don’t know much about San Diego, so it’s really exciting to see the Sea World there. Like, wow... I had no idea it was there. I really want to go. That would just be fascinating. You know. I heard about a lot about the zoo. I just really didn’t think about it till I saw the site. I’m like, oh, yeah, that’s good. You know they have a Hardrock café, so that’s really made my day...”
Phase III: Findings

Proposed model is not supported

– Lower congruence leads to higher satisfaction
– Internet experience leads to lower satisfaction
– Functional needs and hedonic needs
– Novel and exciting information needed
Overall Conclusions

1. Structure
   • Episode and chapter structure
   • Information overloading and decision aid tools
   • Geographical information and time frame

2. Global level of analysis
   • Different languages
   • Marketing language vs. subjective and experiential language

3. Satisfaction
   • High levels of satisfaction
   • Two determinants of satisfaction: functional vs. hedonic, hygiene and motivator
   • Looking for novel and exciting information
Theoretical Implications

– General information search protocol
– Confirmation of dynamic and contingent travel planning process; information overloading; information foraging behavior
– Different levels of behavior and their implications
– Higher level of usability and direction switch in research in use of technology
Design and Managerial Implications

- Based on commonalities of information search behavior
- Apparent limitations: geographical information and time frame
- Beyond customization and personalization
- Collaborations between different parties in tourism
- Generating tourism ontology from consumer’s side
Limitations

- Sampling
- Semantic network analysis
- Choice of destination
- Coding
Future Research

- Large sample for validation
- Different levels of semantic analysis
- More controlled experiment for measuring efficiency
- Analysis on different sectors of tourism
- Analysis of sequence of navigation
Thank You!
Information Seeking and the Internet

- Information search on the Internet as navigation through hypertext
- Navigational task vs. information task (Kim and Hirtle, 1995)
- Information search experience and domain knowledge contribute to successful and satisfactory information search (Hsieh-Yee, 2001)
Why top 25 keywords?

The 25th Keyword “Music”

Figure 4-8. Top 160 Keywords and Their Frequencies in Travelers’ Mental Model
Why top 25 keywords?

Figure 4-6. Top 160 Keywords and Their Frequencies in Tourism Information Space
Information Foraging (Pirolli & Card, 1999)

- Information searchers use proximal cues to identify important information for further exploration or consumption

Information Scent (Chi, Pirolli, Chen and Pitkow, 2001)

- Information searchers identify valuable information from the “snippets” of proximal cues
Comparison of Two Semantic Models, one set with high levels of satisfaction. The other one with low level of satisfaction.

Each individual is unique in terms of their mental models.

Measure of efficiency, maybe the ratio of clicks to episodes of decisions.

Linkage between navigation and comparison of semantics.

Michael: Why they go back to Radisson Hotel web page?

Definition of information hubs… go back and visit.

Qualitative, why they are not satisfied? Why they are satisfied.

Frequencies of keywords, including the distribution of all the keywords.
Research Procedure

1. **Pilot study** (May - June, 2002)
   - 5 subjects

2. **Formal experiment** (September, 2002)
   - 15 diverse subjects (Eveland, & Dunwoody, 2000)
Main Goals of Research

- Understand the structure of travel information search on the Internet
  - Episode?

- Examine satisfaction of travel information search process
  - Discover the congruence two semantic models
  - Explore the relationship between congruence of two models, satisfaction and information searchers’ individual characteristics
Research Procedure

I. Pre-exercise
- Survey
  - Travel Experience
  - Computer and Internet Use Experience
  - Internet Use Experience As Travel Information Source

II. Travel Planning Exercise

III. Satisfaction Survey

IV. Post-experiment Interview

Phase I: Analysis of the process
- SNA on transcripts of interviews
- Travel experience, computer and Internet use experience, the experience of using Internet as travel information source

Phase II: Comparison of two semantic models
- Travel Experience
- Process Satisfaction
- Outcome Satisfaction
- Travel Information Searcher's Model
- Tourism Information Space's Model

Phase III: Modeling Satisfaction
- Internet Use Experience As Travel Information Source
- SNA on web pages
- Reasons for satisfaction