Developing SMART-Search:
A Search Engine to Support the Long Tail in Destination Marketing

Zheng Xiang
School of Merchandising and Hospitality Management
University of North Texas
P.O. Box 311100
Denton, Texas 76203
Tel: 940-369-7680
Fax: 940-565-4348
Email: philxz@unt.edu

Bing Pan
Office of Tourism Analysis
Department of Hospitality and Tourism Management
School of Business and Economics
College of Charleston
Tel: 843-953-2025
Fax: 843-953-5697
Email: panb@cofc.edu

Daniel R. Fesenmaier
National Laboratory for Tourism & eCommerce
School of Tourism and Hospitality Management
Temple University
1700 N. Broad Street
Philadelphia, PA 19122
Phone: 215-204-5611
Fax: 215-204-1455
Email: drfez@temple.edu

TTRA 2008
Submitted for consideration first for oral presentation and then illustrated paper
Category: Innovative marketing approaches

Corresponding Author:
Zheng Xiang
School of Merchandising and Hospitality Management
University of North Texas
P.O. Box 311100
Denton, Texas 76203
Tel: 940-369-7680
Fax: 940-565-4348
Email: philxz@unt.edu
Developing SMART-Search: 
A Search Engine to Support the Long Tail in Destination Marketing

INTRODUCTION

The tremendous growth of the Internet has fundamentally reshaped the way the tourism product is distributed and the way people plan for and consume travel. With the huge amount of information available, search has become the dominant mode in the traveler’s interaction with the Internet for travel planning purposes (TIA, 2005). However, effectively organizing information over the Internet to support travel information search and planning remains a challenging problem (Pan & Fesenmaier, 2006; Werthner & Klein, 1999). Particularly, the increasing dominance of general purpose search engines like Google, Yahoo!, and MSN has led to the low visibility of “the long tail” in tourism i.e., the unique and less well-known tourism enterprises that reflect the idiosyncratic and multi-faceted nature of the destination (Xiang et al., 2007). As such, the goal of this research is to develop an innovative approach to support travel planning on the Internet by addressing the challenges in making “the long tail” available to potential visitors. This paper first conceptualizes “the long tail” within a travel planning context. It then describes the development of a tourism search engine that specifically addresses the challenges in the identification, organization, and representation of “the long tail”. Finally, the implications of this research for destination marketing are discussed.

BACKGROUND

The phrase “the Long Tail” was first coined by Chris Anderson to describe certain technology-enabled business and economic models such as Amazon.com and NetFlix. Those sites have the distribution power to market and sell a greater volume of otherwise hard-to-find items at small volumes than of popular items at large volumes (Anderson, 2006). The aggregated amount of those small niches can actually surpassed those popular products. The notion of “the long tail” can be found in many business domains and is particularly applicable to tourism. That is, the tourism product consists of numerous industry facets (attractions, hotels, restaurants, and places, etc) and the travel experience is comprised of many mental and physical activities that result from the traveler’s interaction with these industry facets (Smith, 1994; Woodside & Dubelaar, 2002). Many, if not most, travelers are looking for unique and idiosyncratic travel experiences and avoid popular and over-crowded ones. Within a trip planning context, a traveler may consider various kinds of unique information about a specific destination relevant to his/her trip; as such, “the long tail” is represented by the idiosyncratic business and non-business entities that, while not necessarily popular, are potentially conducive to planning for trip experiences. A recent study found that the online tourism domain as represented through general purpose search engines like Google, indeed, consists of potentially thousands of entities (Xiang, Wöber, & Fesenmaier, 2007). However, due to the inherent limitations of general purpose search engines and the dominance of certain travel intermediaries (e.g., Expedia), the visibility of “the long tail” for a specific destination is extremely low.

From a destination marketing standpoint, it is argued that the notion of “the long tail” poses a very important perspective. That is, it is important for a DMO (Destination Marketing Organization) to provide the means by which travelers can easily access information about the
domain is idiosyncratic (e.g., small, unique and interesting) tourism-related enterprise that are not the primary attractions of the destination (Gretzel et al., 2006). It is argued that when exposed to the places included in “the long tail”, it is more likely for a traveler to expand his/her consideration set of the trip elements, leading to potentially more opportunities for marketing and promotion (Fesenmaier et al., 2006).

It is not an easy task to develop search technologies that support the user’s interaction with “the long tail” (Wöber, 2006). Specific challenges arise from the identification, organization, and representation of the tourism domain (e.g., Gretzel & Fesenmaier, 2002; Werthner & Ricci, 2004). For example, what constitutes the online domain of tourism is both a theoretical and technical issue (Xiang, et al, 2007). An important challenge arises from the structural representation of the experiential nature of the tourism product with the difficulties in identifying appropriate interface metaphors for representation (Gretzel & Fesenmaier, 2002; Xiang & Fesenmaier, 2005). Also, the studies by Pan and Fesenmaier (2006) and Xiang et al (2007) indicate that there is a substantial difference between how the industry seeks to represent itself online as compared to how tourists conceptualize and describe their experiences. Particularly, the language used by the industry is laden with functional, utility-based vocabularies with the intention for persuasion, while the traveler’s language is subjective, action-based reflecting the quest for personal experiences. This language discrepancy between the representation of the domain and search leads to the lack of effective communication during the human-computer interaction. Therefore, a search technology that provides access to “the long tail” must be able to understand the unique travel desires and needs of an individual user, relate these to the tourism enterprises located in the community, and identify both the related “main attractions” as well as those “unique” places that are consistent with the experiential goals of the user.

SYSTEM DESIGN

The goal of this research is to develop an Internet-based system, referred to as SMART-Search, that can support access to “the long tail” for trip planning purposes. A methodology was employed to ensure the system can represent the tourism domain in a comprehensive and meaningful way so that it provides an enhanced search experience for online travelers. As shown in Figure 1, the system architecture is comprised of eight key components which can be roughly grouped into three sequential developmental processes, i.e., “Domain Identification and Structuring”, “Information Organization”, and “Information Representation.” The current project of “SMART-Search” is based on Chicago as a travel destination. The following describes each of these components.

**Domain Identification and Structuring** follows the method devised in the study conducted by Xiang et al (2007) with the goal to identify Websites that can truly represent a destination. The findings of the study show that, while there are potentially a huge number of relevant Web pages, the tourism domain can actually be represented by a relatively small number of Web sites that contain “the long tail” of tourism. In this stage, expert knowledge was utilized to distinguish tourism related Websites from non-tourism sites, as well as to categorize tourism related Websites into several clusters including informediaries such as CVB websites, intermediaries such as Expedia, industry Websites including individual business and non-business entities, as well as social media sites such as YouTube and Flickr. A compiled list of
Website URLs from this analysis was fed into a Web crawler program, which extracts textual information from these Websites including the title, the meta tags, and the textual content. In addition, thumbnails of the website homepages and certain social media contents were also collected.

![System Architecture of "SMART-Search"

**Information Organization** focuses on effectively organizing and managing the domain data collected in the previous step. The goal is to establish a structure that allows the data fragments (e.g., URLs, text contents, meta tags, and social media contents) to be related in a meaningful way to support the representation of the domain in the user interface. Semantic Web technology (Berners-Lee et al., 2001) is extensively applied in this stage. A set of text mining programs were used to extract a domain ontology that defines the semantic structure of the data. In addition, domain expert knowledge was applied to generate tags that describe non-textual data in social media contents.

**Information Representation** facilitates user experiences when planning trips. System components include the user interface, the query processor, and the recommendation modules. As shown in Figure 2, the user interface utilizes a set of design components that integrate and represent the domain data with the goal not only to provide the relevant results but also to inspire travelers. Specifically, the search results column (in the center) contains three sections including recommended links (e.g., the CVB website), organic links (tourism industry websites), and tourism aggregators (e.g., CitySearch.com). The left column under the search box provides additional suggested keywords for the user to expand the consideration set. This is based on the domain ontology developed in the previous stage. To enrich the search experience, social media
content relevant to a specific query, including pictures and stories, is also included in the right-hand column of the page.

**IMPLICATIONS**

With the growing importance of search in travelers’ use of the Internet, the online distribution of information is increasingly dominated by the “Goliaths” such as general search engines like Google, Yahoo!, and MSN. Unfortunately, these technologies often fail to represent the “richness” of experiences offered at a destination due to the low visibility of small tourism enterprises and the dominance of travel intermediaries. In response, it is argued that tourism domain specific search engines are needed to assist potential travelers to identify both the main attractions (the dominant websites) in a destination as well as those enterprises that offer unique experiences. As such, search technologies developed specifically for the tourism industry, as illustrated in this paper, can have the potentials to allow the industry to form an army of “Davids” to address the challenges in the dynamic market. This research offers several practical implications for the tourism industry and destination marketing on the Internet. Perhaps most important, the capabilities of the proposed search engine make “the long tail” available to potential visitors, enabling destination marketing organizations to further develop their marketing programs and strategies based on collaborative efforts with a large number of local industry partners.
REFERENCES


