



Role of social media in online travel information search

Zheng Xiang^{a,*}, Ulrike Gretzel^{b,1}

^a School of Merchandising and Hospitality Management, University of North Texas, 1155 Union Circle, #311100, Denton, TX 76203-5017, USA

^b Department of Recreation, Park & Tourism Sciences, Laboratory for Intelligent Systems in Tourism, Texas A&M University, 2261 TAMU, College Station, TX 77843-2261, USA

ARTICLE INFO

Article history:

Received 8 August 2008

Accepted 24 February 2009

Keywords:

Social media

Search engine

Travel information search

Online tourism domain

Internet marketing

ABSTRACT

Social media are playing an increasingly important role as information sources for travelers. The goal of this study is to investigate the extent to which social media appear in search engine results in the context of travel-related searches. The study employed a research design that simulates a traveler's use of a search engine for travel planning by using a set of pre-defined keywords in combination with nine U.S. tourist destination names. The analysis of the search results showed that social media constitute a substantial part of the search results, indicating that search engines likely direct travelers to social media sites. This study confirms the growing importance of social media in the online tourism domain. It also provides evidence for challenges faced by traditional providers of travel-related information. Implications for tourism marketers in terms of online marketing strategies are discussed.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

The Internet has fundamentally reshaped the way tourism-related information is distributed and the way people plan for and consume travel (Buhalis & Law, 2008). In recent years, two “mega trends” have noticeably emerged on the Internet, underscoring changes that can significantly impact the tourism system.

On one hand, so-called social media Websites, representing various forms of consumer-generated content (CGC) such as blogs, virtual communities, wikis, social networks, collaborative tagging, and media files shared on sites like YouTube and Flickr, have gained substantial popularity in online travelers' use of the Internet (Gretzel, 2006; Pan, MacLaurin, & Crotts, 2007). Many of these social media Websites assist consumers in posting and sharing their travel-related comments, opinions, and personal experiences, which then serve as information for others. This supports the argument by Thomas Friedman (2006) that “the world is flat”, with consumers gaining substantially more power in determining the production and distribution of information due to the flattening of access on the Internet. At the same time, the Internet also increasingly mediates tourism experiences as tourists use these social media sites to portray, reconstruct and relive their trips (Pudliner, 2007; Tussyadiah & Fesenmaier, 2009).

On the other hand, due to the huge amount of information available, searching has become an increasingly dominant mode in travelers' use of the Internet. For example, research by the Travel

Industry Association of America found that about two-thirds (64%) of online travelers use search engines for travel planning (TIA, 2005). Another recent study showed that search engines serve as the number one online information source for American families in the context of vacation planning (eMarketer, 2008). A series of reports by Internet research firm Hitwise have documented the significance of search engines in terms of generating upstream traffic to tourism Websites (e.g., Hopkins, 2008; Prescott, 2006). As such, search engines have become a powerful interface that serves as the “gateway” to travel-related information as well as an important marketing channel through which destinations and tourism enterprises can reach and persuade potential visitors (Google, 2006; Xiang, Wöber, & Fesenmaier, 2008).

Tourism is an information-intensive industry (Sheldon, 1997; Werthner & Klein, 1999); therefore, it is critical to understand changes in technologies and consumer behavior that impact the distribution and accessibility of travel-related information. Particularly, it has been argued that understanding the nature of the online tourism domain, i.e., the composition of online tourism-related information potentially available to travelers, provides an important stepping-stone for the development of successful marketing programs and better information systems in tourism (Fesenmaier, Wöber, & Werthner, 2006; Xiang et al., 2008). It seems that while social media are, anecdotally, becoming increasingly important in the online tourism domain, there is a lack of empirical data to describe and explain the role of social media in the context of online travel information search.

One of the major questions to be answered concerns the likelihood with which a traveler will be exposed to social media Websites as opposed to other tourism Websites when using

* Corresponding author. Tel.: +1 940 369 7680; fax: +1 940 565 4348.

E-mail addresses: philxz@unt.edu (Z. Xiang), ugretzel@tamu.edu (U. Gretzel).

¹ Tel.: +1 979 862 4043; fax: +1 979 845 0446.

a search engine to find relevant travel information. Therefore, this study aims at investigating the representation of social media as part of the online tourism domain in the context of travel information search supported by search engines. It is hoped that this study can document the current trends on the Internet in order to provide useful insights for online tourism marketing. The rest of the paper is organized as follows: the next section provides the background of this research by critically reviewing existing research on the online tourism domain and social media to provide the rationale for this study. Then, a conceptual framework of travelers' interactions with the online tourism domain through a search engine is presented and research questions are formulated to guide the investigation. In the methodology section, the design of the research using a search engine mining approach is explained. Findings are then presented in correspondence with each of the research questions. Finally, managerial implications for online tourism marketing as well as limitations of this study and plans for future research are discussed.

2. Research background

With the enormous amount of information potentially available to travelers, the Internet constitutes an important platform for information exchange between the consumer and industry suppliers (e.g., hotels, transportation sectors, attractions), intermediaries (e.g., travel agents), controllers (e.g., governments and administrative bodies), as well as many non-profit organizations such as destination marketing organizations (Werthner & Klein, 1999). Different technological interfaces, such as search engines, online travel booking sites, and Websites of destination marketing organizations facilitate the information exchange between online travelers and the so-called "online tourism domain" (Xiang et al., 2008). Presumably, social media, which facilitate the interactions between online consumers, have emerged as an important component of this domain. This section critically reviews the literature on the online tourism domain and social media in tourism, and identifies the limitations of existing literature to provide the rationale for this study.

2.1. Online tourism domain

A domain can be defined as a collection of all informational entities about a specific subject (Hjorland & Albrechtsen, 1995). In the context of the Internet, a domain is the collection of links, domain names, and Web pages that contain texts, images, and audio/video files stored in hypertext formats. The online tourism domain can, therefore, be understood as comprising all such informational entities that relate to travel (Xiang et al., 2008). The study of the online tourism domain can be traced back to Werthner and Klein (1999), who proposed a conceptual framework which delineates the interaction between the consumer and the industry suppliers with the Internet playing a facilitating and mediating role. With the increasing importance of the use of the Internet for travel purposes, more attention has been directed to the analysis of the tourism domain, with an emphasis on the mediating role of specific Internet technologies (e.g., search engines) in representing tourism within a travel planning setting (Pan & Fesenmaier, 2006; Wöber, 2006; Xiang et al., 2008). Pan and Fesenmaier (2006), for example, used the term "online tourism information space" to describe the collection of hypertextual content available for travel information searchers. Wöber (2006) examined one aspect of the tourism domain, i.e., the visibility of tourism enterprises, particularly destination marketing organizations and individual hotel operations in Europe, among six popular search engines. His findings showed that many tourism websites suffer from very low rankings

among the search results, which makes it extremely difficult for online travelers to directly access individual tourism websites through these search engines.

Recently, Xiang et al. (2008) conceptualized the online tourism domain based upon an integration of a number of theoretical perspectives, including: (1) the *industry perspective* (Leiper, 1979, 2008; Smith, 1994), which focuses on what constitutes the supply of tourism and, thus, the organizational entities that comprise the online tourism domain; (2) the *symbolic representation perspective* (Cohen & Cooper, 1986; Dann, 1997; Leiper, 1990), which describes the representation of tourism products and related experiences provided by the industry in various forms; (3) the *travel behavior perspective* (Crompton, 1992; Pearce, 1982; Woodside & Dubelaar, 2002), which includes the activities and the supporting systems at different stages of the travel experience; and, (4) the *travel information search perspective* (e.g., Fodness & Murray, 1998; Gursoy & McLeary, 2004; Vogt & Fesenmaier, 1998), which is related to the nature of the information sought to support travel experiences. Xiang et al.'s (2008) analysis of the domain showed that the representation of tourism through one of the most important interfaces, i.e., a search engine, is extremely rich, reflecting the idiosyncratic nature of destinations and travelers' heterogeneous information needs. In addition, their study demonstrated that popular search engines (e.g., Google) dictate the representation of the domain and, thus, create potential problems for online travelers and tourism suppliers.

2.2. Social media on the Internet

While there is a lack of a formal definition, "social media" can be generally understood as Internet-based applications that carry consumer-generated content which encompasses "media impressions created by consumers, typically informed by relevant experience, and archived or shared online for easy access by other impressionable consumers" (Blackshaw, 2006). This includes a variety of applications in the technical sense which allow consumers to "post", "tag", "digg", or "blog", and so forth, on the Internet. The contents generated by these social media include a variety of new and emerging sources of online information that are created, initiated, circulated, and used by consumers with the intent of educating each other about products, brands, services and issues (Blackshaw & Nazzaro, 2006). In contrast, to content provided by marketers and suppliers, social media are produced by consumers to be shared among themselves. Since more and more travelers seem to tap into this "collective intelligence" available on the Web (Litvin, Goldsmith, & Pan, 2008), this will challenge the established marketing practices of many tourism businesses and destinations.

Social media exist in a variety of forms and serve numerous purposes. Consumer-generated content supported through social media is "a mixture of fact and opinion, impression and sentiment, founded and unfounded tidbits, experiences, and even rumor" (Blackshaw & Nazzaro, 2006: 4). In travel and tourism, past research has focused attention on the socio-psychological aspects of social media use. Not surprisingly, travel-related virtual communities attracted the attention of tourism researchers early on (Kim, Lee, & Hiemstra, 2004; Wang & Fesenmaier, 2003; Wang, Yu, & Fesenmaier, 2002). Virtual tourist communities such as LonelyPlanet and IGoUGo, where tourists can exchange opinions and experiences on topics of common interests, have been around at least since the late 1990s, and several researchers have investigated their roles and impacts in the context of travel.

More recently, new online applications have emerged that add substantially to the information exchange among consumers. Today, Web 2.0, also referred to as "Travel 2.0" in tourism, includes

a range of new technological applications such as media and content syndication, mash-ups, AJAX, tagging, wikis, web forums and message boards, customer ratings and evaluation systems, virtual worlds (e.g., Second Life), podcasting, blogs, and online videos (vlogs) (Schmallegger & Carson, 2008). Consumer blogs have emerged as one of the most prominent themes in research on social media in travel and tourism (Braun-LaTour, Grinley, & Loftus, 2006; Mack, Blose, & Pan, 2008; Pan, MacLaurin et al., 2007; Pudliner, 2007; Pühringer & Taylor, 2008; Waldhör & Rind, 2008). This research underscores the interest in understanding the functions of blogs in creating and sharing new experiences (Pudliner, 2007), its trustworthiness to online travelers (Mack et al., 2008), as well as the use of it as marketing intelligence (Pühringer & Taylor, 2008; Waldhör & Rind, 2008). As evidenced by the success of Websites like tripadvisor.com and zagat.com, online travel-related consumer reviews also represent a significant amount of social media for travel purposes (Gretzel & Yoo, 2008; Vermeulen & Seegers, 2008). The studies on this type of social media focus on its use as well as its impact on travel decision making. Multimedia sharing (i.e., video, photos, podcasting, etc), represented by Websites such as YouTube and Flickr, has attracted tourism researchers by generating interests in understanding the role of this type of social media content in transforming travel experiences (Tussyadiah & Fesenmaier, 2009).

2.3. Research rationale

Understanding the structure and representation of the online tourism domain is important for a better organization of travel-related information by designing effective tourism information systems and for implementing successful marketing campaigns (Werthner & Klein, 1999; Xiang et al., 2008). The structure of the tourism domain has changed drastically owing to the appearance of social media as new players in the field of travel information exchange. However, past research on the online tourism domain has, to a great extent, only considered interactions between the online traveler and the so-called “tourism industry”. As such, it is argued that a comprehensive understanding of the nature of the online tourism domain must appropriately reflect and address the existence, and potential impacts, of social media.

One important limitation in the existing literature on social media is that there is a lack of understanding of its role in online travel information search, especially the likelihood for an online traveler to be exposed to and actually use these social media Websites when looking for travel information. Specifically, currently available information about travelers' use of social media was based upon data collected through self-reported questionnaires (e.g., using a question such as “how often do you use a specific social media Website”) and thus, the degree of objectivity is very limited (e.g., Gretzel & Yoo, 2008). Or, it was based upon controlled experimental settings by asking subjects to conduct a trip planning task online in order to understand the psychological effects of social media on travelers (e.g., Mack et al., 2008). Given these limitations, the extent to which social media constitute the online tourism domain is not well understood in an objective, comprehensive way.

Given the potential impact of social media in online tourism, knowledge about the role of social media in travel information search is considered essential to better inform tourism marketing practices (Fesenmaier, 2007; Gretzel, 2006). First and foremost, the emergence of social media has given rise to issues with respect to how tourism marketers can leverage social media in order to support their online marketing efforts (Gretzel, 2006). Many travel and tourism operations have also recognized the importance of including consumer-generated content on their websites, usually in

the form of edited testimonials (e.g., VisitPA.com and Sheraton.com). Marketing researchers often use the label “electronic word-of-mouth” to describe the impact of such media content (Litvin et al., 2008). Schmallegger and Carson (2008) suggested that strategies of using blogs as an information channel encompass communication, promotion, product distribution, management, and research. It seems that current tourism marketing practice focuses attention on utilizing social media to create positive image and word-of-mouth for tourist destinations and businesses. However, without a solid understanding of the role of social media in online travel information search, tourism marketers' ability to take advantage this “market intelligence” is very limited (Blackshaw & Nazzaro, 2006).

3. Understanding the role of social media in online travel information search

It is argued that the impact of social media in travel and tourism must be understood in relation to the overall online tourism domain and specifically within a travel information search context. By taking into consideration the important role of search engines in travelers' use of the Internet, Fig. 1 provides a conceptual framework illustrating the interactions between an online traveler, a search engine, and the online tourism domain. It is adapted from Xiang et al.'s (2008) original framework, which was based upon a number of past studies focusing on online travel information search behavior (Hwang, Gretzel, Xiang, & Fesenmaier, 2006; Pan & Fesenmaier, 2006; Werthner & Klein, 1999). It includes three key components: (1) the online traveler, who is driven by a number of personal and trip-related needs; (2) the online tourism domain, which is composed of informational entities provided by a number of “players”, including individual consumers through means of social media, this tourism domain has a distinct semantic structure determined by the hypertextual nature of the Internet and the tourism industry structure; and, (3) the search engine, which in large part determines the representation of the tourism domain through the design of interface features, search result rankings, metadata, and paid links and, as a result, influences the traveler's perception and decision making.

This framework is useful in that it stresses the complexity of the online tourism domain and the dominance of search technology. Search engines, to a great extent, define the practical boundaries of the online tourism domain and dictate the way it is represented (Henzinger, 2007; Prescott, 2006; Rose & Levinson, 2004; Spink & Jansen, 2004). Particularly, general-purpose search engines such as Google and Yahoo! have inherently built-in limitations in representing a complex domain like travel and tourism, leading to a series of challenges for destinations and tourism suppliers to provide tourism information to online travelers in a coherent way (Wöber, 2006; Xiang et al., 2008). Search engines can be seen as an information space where businesses in the tourism industry compete for attention of online travelers, because the representation of the domain is largely based upon the ranking and position of search results and will be influenced by contingent factors such as the presence of online advertisements (Kim & Fesenmaier, 2008; Pan, Hembrooke et al., 2007; Spink & Jansen, 2004).

Within this framework, the relationship between search engines and social media is a particularly interesting one. First, social media are updated frequently, which “invites” search engines to index social media pages more frequently. Second, social media, due to their very nature of being socially constructed, usually include a lot of hyperlinks, which will influence their ranking within search results in a positive way. This is particularly due to the fact that social media encourage online consumers to be actively engaged in organizing the contents through activities like “digging” and “tagging”, which, in turn, automatically create an

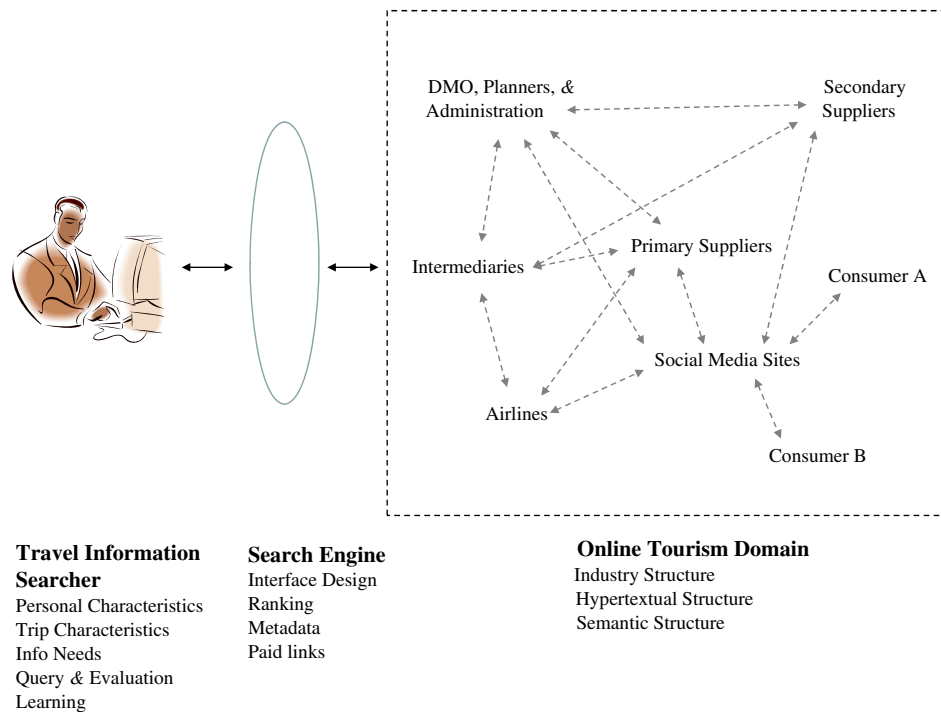


Fig. 1. Social media and tourism domain within the context of using a search engine (Adapted from Xiang, et al., 2008).

ever growing link structure on the Internet. All in all, social media are seen as extremely search engine friendly (Gretzel, 2006). Consequently, online tourism marketers potentially face fierce competition from social media because the search process can lead millions of consumers to highly-relevant social media content pages that can influence attention, awareness, trial, and loyalty levels (Blackshaw & Nazzaro, 2006). As marketers strive to know how to provide attractive, persuasive, and technologically sustainable marketing programs online, they must also compete with consumer-generated content in numerous social media Websites. This is because the very presence of social media potentially erodes the audience pool, impact, as well as reach of their Web-based marketing programs, and can have significant effect on their branding efforts.

Following from the conceptual framework, this study focuses attention on the role of social media in the online tourism domain as represented by a search engine within a travel planning context. Two sets of research questions were formulated, with the first one primarily aiming to understand the extent to which social media are represented through search engines:

Q1a. To what extent does a search engine represent social media Websites in the search results when an online traveler uses the search engine for travel-related information search?

Q1b. What types of social media Websites appear as a result of travel-related information search?

In order to further substantiate the potential impact of social media on travel information search, the second set of questions focuses attention on more specific issues related to search, including:

Q2a. How are social media distributed across search results pages?

Q2b. Do destinations differ from each other in terms of social media results generated? For example, are larger destinations more likely to generate more social media search results than smaller destinations, or vice versa?

Q2c. Are certain types of social media Websites more likely associated with certain keywords used by travelers?

4. Methodology and research design

In order to answer the above research questions, a data mining exercise was devised. A set of keywords were defined in combination with a selected group of destinations in the United States to query a search engine. Content analysis and multivariate analysis approaches were used to understand the data.

4.1. Research design

A trip planning scenario was created by mimicking travelers' use of a search engine when searching for destination-related information. The idea was to examine several aspects of social media as represented by a search engine based on certain queries. These aspects included: (1) the proportion of social media among all the search results retrieved by the search engine; (2) the way the search engine represented social media Websites across different search result pages; (3) the types of social media Websites; and, (4) the relationships between the types of social media Websites and specific search keywords and destinations.

Specifically, a set of 10 pre-defined keywords in combination with nine destination names were used to form queries to search Google. These keywords, including "accommodation", "hotel", "activities", "attractions", "park", "events", "tourism", "restaurant", "shopping", and "nightlife", represent the top level travel-related terms that will likely be used by travelers when they are looking for tourism-related information about a specific destination. The selection of these keywords was based upon a number of past studies and was intended to reflect the "generic" terms or "broad" categories that represent tourism (Pan, Litvin, & O'Donnell, 2007; Wöber, 2006; Xiang et al., 2008).

The focus of the study was on urban destinations to keep the information search context constant. Nine U. S. cities, ranging from large to small in terms of volume of visitation, population size as well as reflecting a certain geographic diversity, were selected to represent urban tourism destinations in the United States. These destinations include New York City, Chicago, Las Vegas, Dallas, Charlotte (NC), San Jose (CA), Elkhart (IN), Bradenton (FL), and Pueblo (CO). This selection was deemed appropriate given the exploratory nature of the study. To avoid confusion with other cities with the same names in different states (e.g., Charlotte, MI), the state abbreviations were appended to the keywords for the less well-known destinations to specify and narrow down the geographic domain of the search queries (e.g., “Charlotte, NC restaurants”).

The search engine provided by Google was used as the technology in focus. Google was chosen because it represents the current state-of-the-art online search technology, and it is highly popular among Internet users and dominant in the online search market. Currently, Google serves the largest percentage of queries (approximately 47.3%) on the Internet, with an index of around 25 billion Web pages and 250 million queries a day (Bertolucci, 2007; Burns, 2007). According to a recent Hitwise (2008) report, Google is even more dominant in the United States by receiving approximately two-thirds of all online queries. Specifically in the travel sector, Google is among the top 10 Websites that generate most of the upstream traffic to travel-specific Websites (Hopkins, 2008). In addition, Google is the most popular text-based search engine on the Internet, which requires the user to type queries into a text search box (Bertolucci, 2007). As such, Google was considered the best “candidate” when assessing aspects related to the nature of the online tourism domain (Xiang et al., 2008).

The 10 keywords in combination with the nine destination names, as mentioned above, resulted in 90 queries, which were entered into Google to obtain the search results. According to the past literature (e.g., Spink & Jansen, 2004), the majority of search engine users will only review search results in the first three pages (assuming 10 search results on one page by default). In this study, search results on the first 10 pages were retrieved in order to provide a more comprehensive representation of social media in the context of online trip planning.

A self-written crawler program designed in Perl programming language was used to automate the process. First, the crawler sent out a query (e.g., “New York City restaurant”) to Google. By specifying the number of pages (10 in this case), the program parsed the content of the search results pages to obtain the URLs associated with each of the “organic” search results. Fig. 2 shows a typical example of organic Google search results. In this case, the URL associated with the first line, i.e., “New York City...”, was extracted and saved into a database along with the destination name and search keywords that were used to generate the search results. The program iterated through this process until all keywords and destination names were used. Google, by default, will provide 10 search results on one single search results page; however, the search engine currently provides a number of additional results on the first search result page for certain queries. Table 1 shows the number of search results extracted for the nine destinations with a total of 10,383. On average, there were 115 search results extracted per query for each destination.

Table 1
Search results retrieved from Google

Destination	Search Results
Chicago	1120
Las Vegas	1121
New York City	1142
Dallas	1160
Charlotte, NC	1180
San Jose, CA	1160
Elkhart, IN	1160
Bradenton, FL	1160
Pueblo, CO	1180
Total	10,383

4.2. Coding and data analysis

Two human coders were used to categorize the Websites included in the search results in a two-step process. First, the coders were instructed to categorize the 10,383 search results into social media and non-social media Websites. The coders accessed the Web pages by following the URLs and indicated whether the Website in which the Web page was situated was a social media Website or not. Intercoder reliability was checked using Krippendorff's Alpha (Krippendorff, 2004), which was 0.82, indicating there was a high level of agreement between the two coders. In total, 1150 search results were identified as social media Websites, which constituted approximately 11% of all the search results extracted.

The second step involved coding the identified Websites into different types of social media. While there is no existing typology for social media on the Internet, five main categories were created by the authors after an extensive review of existing literature. They included: virtual community sites such as LonelyPlanet and IgoUgo, consumer review sites such as TripAdvisor and InsiderPages, personal blogs and blog aggregators such as blogspot.com, social networking sites such as Facebook and MySpace, and media sharing sites such as YouTube and Flickr. A sixth category, namely “other”, was created in case there were websites that did not necessarily fall into these five categories. After the coding was completed, intercoder reliability was once again checked. This time, the Krippendorff's Alpha was 0.74, which was a bit lower than the first round of coding. Considering that it involved more categories, the results of the second round of coding still indicate reasonably good intercoder reliability. The differences in coding were resolved by the researcher discussing with the coders until an agreement was reached for each specific case.

In order to answer the research questions, the data analysis involved two steps. First, descriptive analyses were conducted to describe social media represented in Google by identifying the unique domain names (Websites) for social media. In the second step, several aspects of the representation of social media were examined, including: (1) the distribution of social media Websites across the top 10 search results pages in Google; (2) the distribution of social media Websites by keywords; and, (3) the distribution of social media Websites by destinations. Finally, a correspondence analysis was conducted in order to explore the relationship between keywords and types of social media. This was intended to answer the question whether certain keywords would more likely

New York City.com: New York Restaurants and Dining Guide Manhattan ...
New York City.com Guide to **New York Restaurants** and Manhattan **Restaurants**.
 NYC.com, the authentic **city** site offers a complete dining guide to all **New York ...**
www.nyc.com/restaurants/ - 114k - [Cached](#) - [Similar pages](#) - [Note this](#)

Fig. 2. A typical Google search result.

generate certain types of social media in search results (e.g., social networking vs. virtual communities). A proximity matrix was established by using the number (frequencies) of co-occurrences of the type of social media Website and a specific keyword. Then, a correspondence map was generated to show the associations between keywords and social media type.

5. Findings

First, the descriptive analysis results are presented to show the extent to which social media constitute the online tourism domain represented by Google. Specifically, this analysis focused on identifying the top domains (i.e., Websites) representing social media as well as the types of social media included in the search results. The second set of results shows the associations between social media and specific search queries.

5.1. Social media represented by Google

Among the total 10,383 search results, there were 1150 (approximately 11%) identified as search results representing social media. Given the otherwise rather fragmented nature of the search results, this suggests that social media, indeed, represent a substantial part of the online tourism domain and play an important role. Fig. 3 shows the distribution of unique domain names among these social media search results. In total, there were only 335 unique domain names which represented the 1150 search results of social media. In this graph, the top 20 (6%) unique domain names with the highest frequencies represent approximately 50% of all the 1150 Websites; the top 68 (18%) domain names represent approximately 70% of all the social media Websites; and, the top 118 (35%) unique domain names represent 80% of all the social media Websites. As such, there seems to be a “core” and a “long tail” in this distribution. That is, there are a relatively small number of Websites forming the core of the social media portion of the domain, while a significant number of Websites which only occurred once or twice represent the long tail of the domain. In other words, the social media represented by Google are dominated by a handful of “big players”.

Fig. 4 shows the breakdown of the types of social media. Not surprisingly, virtual communities represent approximately 40% (N = 456) of all the 1150 identified social media sites. This seems to indicate the status of virtual communities, i.e., Websites such as igougo.com and lonelyplanet.com, as the primary platforms for

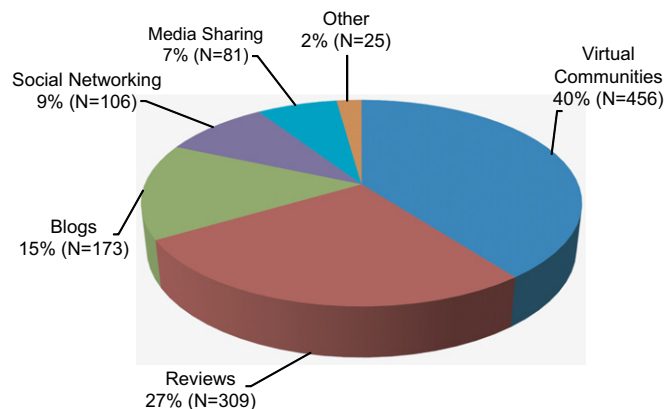


Fig. 4. Composition of social media represented in Google.

online travelers to share their experiences. The longer history of these sites might also play a role. Consumer review sites (e.g., zagat.com and tripadvisor.com) come next, representing about 27% (N = 309) of all social media sites. Personal blogs and blog aggregators (e.g., blogspot.com) also represent a substantial part of social media (15%; N = 173) retrieved by Google in the context of travel information search. These three categories, in combination, constitute over 80% of social media represented in Google. In addition, there are social networking sites (e.g., facebook.com) and media sharing sites (e.g., youtube.com), representing 9% (N = 106) and 7% (N = 81) of all social media sites, respectively. The last type, namely “other”, includes sites that could not be easily categorized (e.g., Wikipedia-type sites), constituting 2% of all social media sites.

The top 20 social media Websites are shown in Table 2. As can be seen, these Websites represent the most “popular” (as determined by Google) social media Websites that contain travel-related content. Interestingly, this top 20 list represents several types of social media Websites. For example, it includes virtual community sites such as virtualltourist.com, and igougo.com; consumer review sites such as yelp.com, travelpost.com, and zagat.com; personal blog sites such as blogspot.com; and, social networking tools such as meetup.com and myspace.com. However, it seems that other types of social media such as video/audio media sharing sites (such as youtube.com and flickr.com) are not well represented among the most prominently displayed Websites.

Table 2
Top 20 unique domain names among social media search results

Domain	Frequency	Cumulative Percentage
tripadvisor.com	96	8.3%
virtualltourist.com	78	15.1%
igougo.com	58	20.2%
mytravelguide.com	54	24.9%
yelp.com	37	28.1%
meetup.com	30	30.7%
travelpost.com	26	33.0%
insiderpages.com	25	35.1%
associatedcontent.com	17	36.6%
yellowbot.com	17	38.1%
blogspot.com	16	39.5%
myspace.com	16	40.9%
gusto.com	15	42.2%
outside.in	13	43.3%
realtravel.com	13	44.4%
topix.com	13	45.6%
pubcrawler.com	12	46.6%
zagat.com	11	47.6%
fodors.com	10	48.4%
lonelyplanet.com	10	49.3%

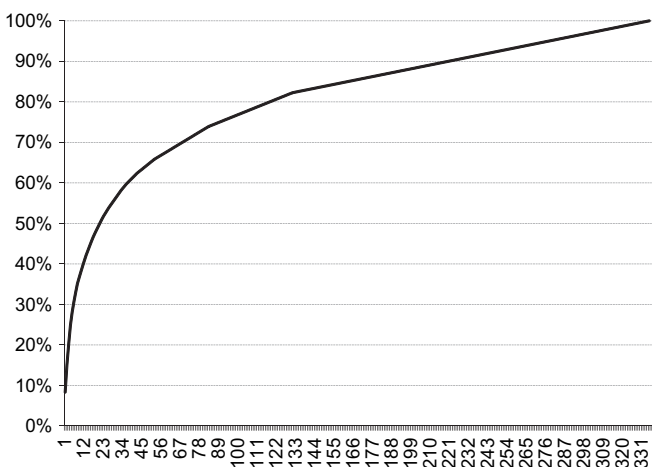


Fig. 3. Distribution of unique domain names among social media websites.

Fig. 5 shows the distribution of the social media Websites across the sampling frame of the study, i.e., the first 10 search results pages in Google. As can be seen, every search results page contains roughly 10% of search results that are social media sites. That is, there is approximately one search result that is a reference to a social media Website on each and every Google search results page. It is quite interesting to observe that social media are homogeneously distributed across Google search results pages.

5.2. Associations between social media and search queries

The goal of this analysis was to examine whether there is any relationship between specific search queries and social media representation. Fig. 6 shows the distribution of social media Websites among the selected destinations. As can be seen, the number of social media Websites for most of the destinations remains very similar except for Elkhart, IN and Bradenton, FL. While this is a small sample of destinations, it seems that the number of social media sites for the “large” destinations are relatively stable, while there might be more variations for relatively “small” destinations.

Fig. 7 shows the distribution of social media sites among the 10 pre-defined keywords. Among these keywords, “nightlife” and “restaurant” generated the most social media-related results (21.5% and 14.3%, respectively). Other keywords also generated a substantial number of social media sites, ranging from 6.4% (attractions) to 10.5% (tourism). This indicates that the sizes of the sub-domains represented by these keywords are quite different. It seems that consumers tend to produce more online contents on certain activities such as nightlife and dining.

Fig. 8 shows the correspondence map of type of social media and keywords used for search. The test of independence between rows and columns in the similarity matrix generated based upon the co-occurrences of types of social media and keywords indicated that the rows and columns were highly associated. The overall inertia was approximately 0.30, indicating it explained about 30% of the total variance in the data. As can be seen from the correspondence map, certain types of social media sites are interestingly grouped together, along with certain types of keywords. For example, virtual community sites are highly associated with the keywords “tourism”, “activities”, “attractions”, “accommodations”, and “hotel”. This seems to suggest that virtual communities play an important role in the sharing of travelers’ experiences related to these “core” tourism activities. Sites for social networking, photo/video sharing, and blogs are closely grouped together and they are associated with keywords such as “events”, “nightlife”, and “park”. This suggests that there is a substantial number of social networking, media sharing, and personal blog sites providing space for travelers to share their experiences of nightlife and events and, thus, allowing many others to plan for these types of experiences.

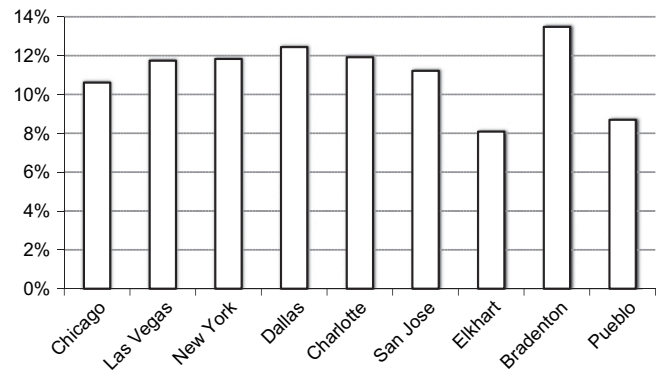


Fig. 6. Distribution of social media by destination.

Not surprisingly, consumer review sites are closely related to keywords such as “shopping” and “restaurants” as well as “hotels”, indicating that there is a relatively large number of Websites that provide opinions about shopping, dining and hotels.

6. Discussion

Given their important role in facilitating travelers’ access to the online tourism domain, search engines directly and indirectly promote social media to travel information searchers. This study explored the extent to which social media Websites are represented in one of the most powerful search engines, i.e., Google. Although the investigation involved a very limited number of destinations and a handful of search queries that are likely used by travelers, the findings revealed interesting dynamics in some of the key areas in online tourism. Thus, this research contributes to the understanding of the online tourism domain and its implications for online tourism marketing in a number of ways.

The findings of the study show that social media, indeed, constitute a substantial part of the online tourism domain and, thus, play an important role within the context of trip planning using a search engine. However, they do not consume all spaces on search results pages and, thus, leave room open for tourism marketers to effectively compete with social media for consumers’ attention. Social media include a variety of Websites that allow consumers to share their experiences in different ways, ranging from posting their stories, their comments, to even their pictures and movie clips. One would presume that, because of their distributed nature and recency, they may not enjoy the same status in Google as other contents provided by industry suppliers. However, social media appear on the first few search results pages

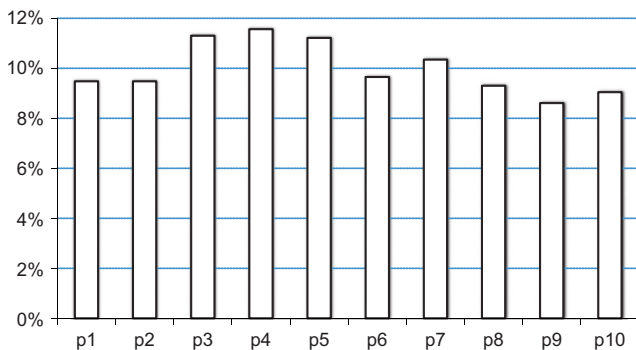


Fig. 5. Distribution of social media across the top 10 search result pages.

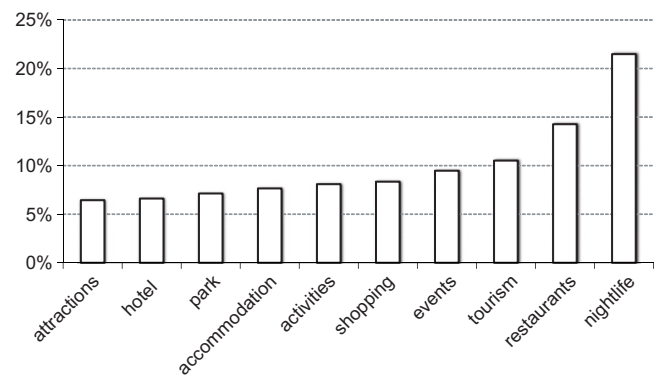


Fig. 7. Distribution of social media by keyword.

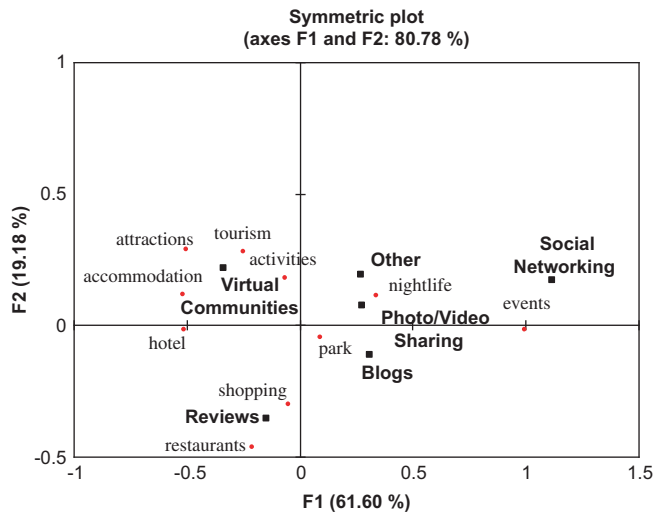


Fig. 8. Correspondence between keywords and types of social media.

in Google, instead of being “buried” somewhere less likely to be accessed by online travelers. This suggests that these social media sites are quite substantial in terms of the size of their sites, the up-to-date nature and relevance of their contents, and the level of connectivity with other sites on the Internet, considering the specific ranking algorithms used by Google. Interestingly, the domain of social media seems to follow the rule of the “long tail” economy (Anderson, 2006) in that a handful of Websites represent a substantial size of the domain, while there are a large number of smaller Websites representing numerous, less popular entities in the domain.

The findings also revealed some interesting relationships between social media and search queries a traveler is likely to use. Specifically, it seems that the role of social media is not dependent on the specific destination in that the domain for most destinations (at least among the relatively small sample used in this study) consists of roughly the same amount of social media Websites that provide consumer-generated contents. Differences were only found for the smaller cities. However, certain keywords (e.g., nightlife and restaurants) are clearly more likely to generate more social media search results as compared to others (e.g., attractions). This indicates that the size and structure of the sub-domains represented by these keywords might be substantially different from each other. In addition, it seems that certain keywords are more likely to generate certain types of social media within a trip planning context. Specifically, virtual community Websites are more closely tied to the “core” tourism businesses such as attractions, activities, and accommodations, while consumer review sites are related to shopping, hotels and restaurants, and, social networking, blogs, and photo/video sharing sites with events, nightlife, and parks.

7. Conclusions and implications

While exploratory in nature, this study offers several useful insights into the online tourism domain, the role of search engines in representing this domain, as well as the implications for tourism marketers. Specifically, this study first confirms the growing importance of social media in the online tourism domain for travel information search. Particularly, while existing literature focuses attention on the socio-psychological effects of social media in travel and tourism (e.g., Pan, MacLaurin et al., 2007; Pudliner, 2007; Pühringer & Taylor, 2008; Wang & Fesenmaier, 2003), very little is known in terms of the extent to which social media actually

constitute the online tourism domain. By showing the extent to which this type of information is made accessible to travelers through one of the most important channels (i.e., search engines), this study fills the gap in the existing tourism literature regarding what, indeed, constitutes online tourism (Werthner & Klein, 1999; Xiang et al., 2008).

Second, this study provides a preliminary understanding of the potential tension between two types of travel information providers, i.e., the tourism industry and online consumers. The findings of the study revealed that social media Websites are “ubiquitous” in online travel information search in that they occur everywhere (i.e., on different search results pages in Google and for different tourist destinations) no matter what search keywords a traveler uses. Certain social media Websites such as TripAdvisor, VirtualTourist, and IgoUGo, which can be considered more comprehensive and travel-specific sites, are becoming increasingly popular and are likely to evolve into primary online travel information sources. In addition, the growth of social media is not only represented by these frequently used Websites but also by the existence of different types of social media and numerous small sites within a travel information search setting. Particularly, blog sites (e.g., travelpost and blogspot) and social networking sites (e.g., meetup) are making inroads into the territories that used to be dominated by traditional suppliers. The results confirm that tourism marketers can no longer ignore the role of social media in distributing travel-related information without risking to become irrelevant.

Finally, this study provides insights into the means by which tourism marketers can tap into the social information space by understanding what drives online travelers to social media sites. Tourism marketers are facing challenges resulting from the shift in distribution channels and the emergence of new media (Fesenmaier, 2007; Werthner & Klein, 1999). In response to these changes, tourism marketers need to understand the technological dynamics in order to better reach out and promote their businesses and destinations to online travelers. With the recent changes on the Internet that allow for easy content generation, consumers are gaining more power over what and how information is distributed and used on the Internet (e.g., Tapscott & Williams, 2006). Search engines constitute such a technology for which intimate knowledge seems to be critical in the context of tourism marketing.

Given their dominant role in online information search, search engines provide the “jump board” for online travelers to access social media sites. Within such a context, search keywords as well as the way search engines present search results (e.g., ranking, metadata, etc) will determine the likelihood a traveler will review and choose to visit a specific Website (Jansen, Brown, & Resnick, 2007). For example, it will be useful for marketers to know what types of search queries will likely generate more search results representing social media content, or, what types of social media emerge as search results for a specific destination. Such knowledge will allow tourism marketers to effectively monitor the shifts in channels of information distribution in order to develop successful strategies for their online marketing efforts. Indeed, the results of this study provide inputs based on which tourism marketers can reflect on their current online marketing strategies.

Since it has been argued that the future of Internet-based tourism will be focused on consumer centric technologies that will support tourism organizations in interacting with their customers dynamically (Buhalis & Law, 2008), this study provides useful insights into the new avenues tourism marketers need to pursue in order to achieve such a goal. The findings clearly indicate a great need for search engine optimization efforts as well as search engine advertising to ensure that a tourism Website is represented and can compete with the rankings of social media sites. An alternative

strategy is to embrace social media and (1) advertise or provide contents on those sites or (2) integrate social media components on the tourism destination or supplier Website. With the so-called Web 2.0 technologies, such integration is not only possible but also relatively easy.

8. Limitations and future research

Given its exploratory nature, this study has several limitations. In addition to the lack of comprehensiveness due to the limited number of keywords and destinations selected, this study employed a cross sectional approach in that the data reflected only a snapshot of the social media represented through a popular search engine at one specific point in time and for a specific group of destinations. Obviously, it is more desirable to conduct longitudinal studies capturing the role of social media over time in a dynamic environment such as the online tourism domain. Also, other mainstream search engines such as Yahoo! and MSN handle a substantial number of online queries and they are popular among online travelers as well (Hopkins, 2008; Prescott, 2006). These search engines should be included in future analyses to reflect the mediation of these technologies in a more comprehensive way. Further, future studies should focus on improving the external validity for this line of research by including more destinations, reflecting a greater range of types of destinations and geographic areas. This will also allow for additional comparisons and analysis of the online tourism domain beyond the context used in this study. A goal of future research could also be the development of practical tools (e.g., benchmarking systems) to keep track of the change in the role of social media in order to provide useful and timely insights for online tourism marketers.

References

- Anderson, C. (2006). *The long tail: Why the future of business is selling less for more*. New York: Hyperion.
- Bertolucci, J. (2007). Search engine shoot-out. *PC World*, 25(6), 86–96.
- Blackshaw, P. (2006). The consumer-generated surveillance culture. <http://www.clickz.com/showPage.html?page=3576076>. Accessed 20.08.07.
- Blackshaw, P., & Nazzaro, M. (2006). *Consumer-generated media (CGM) 101: Word-of-mouth in the age of the web-fortified consumer*. New York: Nielsen BuzzMetrics.
- Braun-LaTour, K. A., Grinley, M. J., & Loftus, E. F. (2006). Tourist memory distortion. *Journal of Travel Research*, 44(4), 360–367.
- Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—the state of eTourism research. *Tourism Management*, 29(4), 609–623.
- Burns, E. (2007). U.S. search engine rankings and top 50 web rankings. <http://searchenginewatch.com/showPage.html?page=3625081>. Accessed June 2007.
- Cohen, E., & Cooper, R. L. (1986). Language and tourism. *Annals of Tourism Research*, 13(4), 533–563.
- Crompton, J. L. (1992). Structure of vacation destination choice sets. *Annals of Tourism Research*, 19(3), 420–434.
- Dann, G. M. S. (1997). *The language of tourism: A sociolinguistic perspective*. Wallingford, UK: CAB International.
- eMarketer. (2008). First summer vacation stop: the Internet. http://www.emarketer.com/Article.aspx?id=1006344&src=article1_newsltr. Accessed 02.05.08.
- Fesenmaier, D. R. (2007). Introduction: challenging destination promotion. *Journal of Travel Research*, 46(1), 3–4.
- Fesenmaier, D. R., Wöber, K., & Werthner, H. (2006). Introduction: recommendation systems in tourism. In D. R. Fesenmaier, K. Wöber, & H. Werthner (Eds.), *Destination recommendation systems: Behavioral foundations and applications*. Wallingford, UK: CAB.
- Fodness, D., & Murray, B. (1998). A typology of tourist information search strategies. *Journal of Travel Research*, 37(2), 108–119.
- Friedman, T. L. (2006). *The world is flat*. New York: Farrar, Straus and Giroux.
- Google. (2006). Seattle's convention and visitors bureau found 30% ROI with Google AdWords. <http://www.google.com/ads/scvb.html>. Accessed 15.12.06.
- Gretzel, U. (2006). Consumer generated content – trends and implications for branding. *e-Review of Tourism Research*, 4(3), 9–11.
- Gretzel, U., & Yoo, K. H. (2008). Use and impact of online travel reviews. In P. O'Connor, W. Höpken, & U. Gretzel (Eds.), *Information and communication technologies in tourism 2008* (pp. 35–46). New York: Springer.
- Gursoy, D., & McLeary, K. W. (2004). An integrated model of tourists' information search behavior. *Annals of Tourism Research*, 31(2), 343–373.
- Henzinger, M. (2007). Search technologies for the Internet. *Science*, 317(5837), 468–471.
- Hitwise. (2008). Google receives 68 percent of U.S. Searches in May 2008. <http://hitwise.com/press-center/hitwiseHS2004/leader-record-growth.php>. Accessed June 2008.
- Hjorland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*, 46(4), 400–425.
- Hopkins, H. (2008). Hitwise US travel trends: how consumer search behavior is changing. Available from. <http://www.hitwise.com/registration-page/hitwise-report-travel-trends.php>.
- Hwang, Y. H., Gretzel, U., Xiang, Z., & Fesenmaier, D. R. (2006). Information search for travel decisions. In D. R. Fesenmaier, K. Wöber, & H. Werthner (Eds.), *Destination recommendation systems: Behavioral foundations and applications* (pp. 3–29). Wallingford, UK: CAB.
- Jansen, B. J., Brown, A., & Resnick, M. (2007). Factors relating to the decision to click on a sponsored link. *Decision Support Systems*, 44(1), 46–59.
- Kim, H., & Fesenmaier, D. R. (2008). Persuasive design of destination Websites: an analysis of first impression. *Journal of Travel Research*, 47(1), 3–13.
- Kim, W. G., Lee, C., & Hiemstra, S. J. (2004). Effects of an online virtual community on customer loyalty and travel product purchases. *Tourism Management*, 25(3), 343–355.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks: SAGE.
- Leiper, N. (1979). The framework of tourism: towards a definition of tourism, tourist, and the tourist industry. *Annals of Tourism Research*, 6(4), 390–407.
- Leiper, N. (1990). Tourist attraction systems. *Annals of Tourism Research*, 17(3), 367–384.
- Leiper, N. (2008). Why 'the tourism industry' is misleading as a generic expression: the case for the plural variation, 'tourism industries'. *Tourism Management*, 29(2), 237–251.
- Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), 458–468.
- Mack, R. W., Blose, J. E., & Pan, B. (2008). Believe it or not: credibility of blogs in tourism. *Journal of Vacation Marketing*, 14(2), 133–144.
- Pan, B., & Fesenmaier, D. R. (2006). Online information search: vacation planning process. *Annals of Tourism Research*, 33(3), 809–832.
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: users' decisions on rank, position and relevancy. *Journal of Computer-Mediated Communication*, 12(3), 801–823.
- Pan, B., Litvin, S. W., & O'Donnell, T. E. (2007). Understanding accommodation search query formulation: the first step in putting 'heads in beds'. *Journal of Vacation Marketing*, 13(4), 371–381.
- Pan, B., MacLaurin, T., & Crofts, J. C. (2007). Travel blogs and their implications for destination marketing. *Journal of Travel Research*, 46(1), 35–45.
- Pearce, P. (1982). *The social psychology of tourist behavior*. Oxford: Pergamon Press.
- Prescott, L. (2006). Hitwise US travel report. Available from. <http://www.hitwise.com/registration-page/hitwise-us-travel-report.php>.
- Pudliner, B. A. (2007). Alternative literature and tourist experience: travel and tourist Weblogs. *Journal of Tourism and Cultural Change*, 5(1), 46–59.
- Pühringer, S., & Taylor, A. (2008). A practitioner's report on blogs as a potential source of destination marketing intelligence. *Journal of Vacation Marketing*, 14(2), 177–187.
- Rose, D. E., & Levinson, D. (2004). Understanding user goals in Web search. Paper presented at the WWW 2004, New York, NY.
- Schmallegger, D., & Carson, D. (2008). Blogs in tourism: changing approaches to information exchange. *Journal of Vacation Marketing*, 14(2), 99–110.
- Sheldon, P. J. (1997). *Tourism information technology*. Oxon, UK: CAB International.
- Smith, S. L. J. (1994). The tourism product. *Annals of Tourism Research*, 21(3), 582–595.
- Spink, A., & Jansen, B. J. (2004). *Web search: Public searching of the Web*. New York: Kluwer.
- Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything*. New York: Penguin Group.
- TIA. (2005). *Travelers' use of the Internet*. Washington, DC: Travel Industry Association of America.
- Tussyadiah, I., & Fesenmaier, D. R. (2009). Mediating tourist experiences: access to places via shared videos. *Annals of Tourism Research*, 36(1), 24–40.
- Vermeulen, I. E., & Seegers, D. (2008). Tried and tested: the impact of online hotel reviews on consumer consideration. *Tourism Management*, 30(1), 123–127.
- Vogt, C. A., & Fesenmaier, D. R. (1998). Expanding the functional information search model. *Annals of Tourism Research*, 25(3), 551–578.
- Waldhör, K., & Rind, A. (2008). eBlogAnalysis – mining virtual communities using statistical and linguistic methods for quality control in tourism. In P. O'Connor, W. Höpken, & U. Gretzel (Eds.), *Information and communication technologies in tourism 2008* (pp. 453–462). New York: Springer.
- Wang, Y., & Fesenmaier, D. R. (2003). Assessing motivation of contribution in online communities: an empirical investigation of an online travel community. *Electronic Markets*, 13(1), 33–45.
- Wang, Y., Yu, Q., & Fesenmaier, D. R. (2002). Defining the virtual tourist community: implications for tourism marketing. *Tourism Management*, 23(4), 407–417.

- Werthner, H., & Klein, S. (1999). *Information technology and tourism: A challenging relationship*. Vienna: Springer.
- Wöber, K. (2006). Domain specific search engines. In D. R. Fesenmaier, K. Wöber, & H. Werthner (Eds.), *Destination recommendation systems: Behavioral foundations and applications*. Wallingford, UK: CABI.
- Woodside, A. G., & Dubelaar, C. (2002). A general theory of tourism consumption systems: a conceptual framework and an empirical exploration. *Journal of Travel Research*, 41(2), 120–132.
- Xiang, Z., Wöber, K., & Fesenmaier, D. R. (2008). Representation of the online tourism domain in search engines. *Journal of Travel Research*, 47(2), 137–150.