



Journal of Internet Banking and Commerce

An open access Internet journal (<http://www.arraydev.com/commerce/jibc/>)

*Journal of Internet Banking and Commerce, April 2013, vol. 18, no.1
(<http://www.arraydev.com/commerce/jibc/>)*

A Study of the Exposure of Hotels Information on the Internet: A Case Study of Sabah Hotels of Malaysia

ZAKARIYA BELKHAMZA, PhD

Senior Lecturer, School of Business and Economics, Universiti Malaysia Sabah, Malaysia

Postal Address: School of Business and Economics, Universiti Malaysia Sabah, Jalan UMS, Teluk Likas, 88450, Kota Kinabalu, Sabah, Malaysia.

Email: zakariya@ums.edu.my

Dr. Zakariya Belkhamza is a Senior Lecturer at the School of Business and Economics, Universiti Malaysia Sabah, where he teaches Management Information Systems subjects at both undergraduate and postgraduate level. He holds a PhD in Management and a Master's degree in Business. Beside his teaching experience, he is also involved in many research and consultancy activities in Malaysia, Thailand, Turkey, and Algeria. His research interests include IT management, strategic Information Systems, IS implementation, and IS assessment and evaluation. His research appeared in many international journals and conference proceedings.

Abstract

The Internet nowadays has become an important and major information search channels for travelers looking for accommodation and services provided by hotels. Hotels' managers have adopted the Internet as one of their primary marketing channels to attract customers while website development has become one of their business infrastructures for online bookings and transactions. On the other hand, hotels' websites are the main search target for travelers. Especially for hoteliers, owning a website is an important step in order to increase visibility and brand familiarity. However, most hotels are not aware of the important of making their websites visible to search engines and the nature of their presence on the Internet. The objective of this paper is to provide an actual assessment and evaluation on the current websites of Sabah hotels in Malaysia and analyze the web information richness in the search engines. The paper further

examines the search engines index ratio of the websites of Sabah hotels and provides interesting insights and recommendations on how to improve the visibility and presence on the Internet

Keywords: search engine, index ratio, Sabah hotels, hotel website, web content, sitemap, internet exposure.

© Zakariya Belkhamza, 2013

INTRODUCTION

The Internet has become a major information search channels for travelers looking for accommodation and services provided by hotels. Tourism organizations have adopted the Internet as one of their primary marketing channels to attract customers (Buhalis & Law, 2008). Hotel websites are the main search target for travelers. Many companies have invested a huge amount of money in order to create and maintain their websites to represent their business activities. Specifically for hoteliers, owning a website is an important step in order to increase visibility and brand familiarity (Hashim & Murphy, 2007). However, this investment does not necessary mean a good return on investment (Hernández, Jiménez, & Martín, 2009).

The continuous development of powerful search engines had made it easier for customers to use keywords to look for information on the Internet rather than typing the domain name. According to Pew Internet & American Life Project (2012), 91% of web surfers use search engines to find information needed. Although this way of search for information may also benefit the hotel organizations if their information exists in the Internet, it has a negative impact on the name branding when the hotel does not own a website, or when the website is not well placed or found in the search engine database. This challenge is increasing as the number of websites and domain names is dramatically increasing. Amid of this phenomenon, most hotels are not aware of the importance of making their websites visible to search engines and the nature of their presence in the Internet. Having updated contents of the website does not guarantee a top search engine indexing. As Sabah is a major tourism destination in Malaysia, it is important to evaluate the actual hotels' websites position in the search engines database in order to observe the hotels' exposure on the Internet. This is a crucial step for hotels to enhance their visibility and fully utilize the Internet as a marketing channel.

Therefore, the main objective of this paper is to provide an actual assessment and evaluation on the current websites of Sabah hotels and analyze the web information richness in search engine. The paper further examines the search engines index ratio of the websites of Sabah hotels and provides recommendations and suggestions on how to improve the visibility and the presence in the Internet.

LITERATURE REVIEW

Due to the information-based nature of the hospitality business, Internet is becoming an important means of distributing and promoting tourism services (Doolin, Burgess, & Cooper, 2002). Web-based interfaces and websites are the key element of the World Wide Web technology in destination marketing, which indicate a remarkable improvement over the past 15 years (Gretzel, Yuan, & Fesenmaier, 2000 ; Hanna & Millar, 1997). Hoteliers often use the Internet to design websites for ease and convenience, initiate web campaigns, develop brands and integrate branding strategy, widen their marketplace and improve customer services (Mullen, 2000).

In order to maximize the benefits of these advantages from the Internet, hotel organizations should take the utilization of the latest technologies and development in the Internet technologies. With the growing importance of search engine, hoteliers should make sure that the information quality representing them in the search engines databases meets the customers' need and expectations (Leung & Law, 2012). According to Klein (2002) and Knight & Burn (2005), good quality of information provided by search engines must be up-to-date, visible and popular. However, because of the advanced technologies of Internet, the technical functionality of search engines is becoming more complex and sophisticated especially crawlers and robots, in which it has created a competitiveness of visibility among websites. This situation makes it difficult for less aggressive websites to be in the front pages of the search engines when searched by less attracted keywords.

Due to the huge number of websites updated every day, search engines frequently clean and remove older pages, keeping only the most updated records (Lewandowski, 2008). This means that if a websites remains outdated in their web servers, search engines will retain that record in the database for a longer period even if an updated occurred later. The popularity of the website will determine how soon an update would appear in the database. In order to increase the website popularity, website should drive a large amount of traffic. A proper indexing of the website contents in the search engine database is the key factor for a successful search engine marketing (Paraskevas, Katsogridakis, Law, & Buhalis, 2011).

In order for a hotel website to be listed by major search engine, a first step is to register the website's URL in a search engine's directory. Submitting a sitemap will also increase crawling frequency and indexing coverage. This will also ensure that the robot will collect more accurate information and more web contents of the websites.

In Malaysia, very few studies were conducted to evaluate the current position of the websites of hotels in the Internet and to what extent they are visible and exposed (Hashim & Murphy, 2007; Hashim, Murphy, Purchase, & O'Connor, 2010). According to the Malaysian Ministry of Tourism's directory, there are 219 hotels in Sabah. In the recent year, 2011, Sabah had 326 hotels, a drastic increase of 53.1 % from the previous year, the second highest figure after Selangor state. The number of rooms supply in 2011 was 18,078, with 52.5% increase from the previous year, the highest increase in Malaysia. On the other hand, the recent statistics also show that in 2011, the hotel guests in Sabah was 3,547,172, with % 2.57 lesser than the previous year 2010. The average occupancy rate for 2011 was 58.8, representing a 2.6% increase from the last year.

These statistics indicate that although the accommodation infrastructure and hotels and rooms supply in Sabah are promising, there are still much work to be done by hotels' managers in order to achieve the best performance and meet the tourism demand in Sabah.

RESEARCH METHODOLOGY

In order to evaluate the exposure of websites of Sabah hotels in the Internet, 130 hotels names were collected from Sabah Tourism Board. A first screening of the website of each hotel was performed to ensure the technical accuracy of the URL. The definition of a hotel website adopted in this study is the official URL of a commercial establishment providing lodging, meals, and other services for travelers. This analysis resulted to an elimination of 12 URLs of which were not necessarily the official websites of an accommodation service. These websites were mainly personal blogs of lodging owners or travel agencies that promote lodging or homestay services. The final list of hotels' websites used for this study was 108 websites. This indicates that 82% of hotel accommodation businesses in Sabah own an official website.

To analyze the web contents of the hotels' websites, two sets of data were needed. The first set is the number of content pages on the website's sitemap. Sitemap is defined as a visual or textually organized model of a Web site's content that allows the users to navigate through the site to search the information (Rouse, 2005). This data was collected using a sitemap generator tool, where the software scans the whole structure of the hotel's domain name and generates a list of all pages available under that URL. The second set of data is the number of web pages of that URL indexed by the search engine. This was done by recording the number of web pages of each hotel's website appears in the search engine. These two sets of data will enable us to have a clear picture on the visibility and popularity of the website on the Internet. To further investigate this issue, the search engine index ratio for the website was also calculated.

FINDINGS AND DISCUSSION

The 108 hotel websites included in this study were classified into three categories. An international chain, this refers to a hotel chain which has sister hotels in multiple continents. The second category is local chain, which refers to a hotel chain which has sister hotels either in Sabah or other part of Malaysia. The third category, independent hotel refers to any other hotel which is not affiliated with any chain or network. Out of 108 hotels analyzed, 8.3% are affiliated with international chains, 27.8% are affiliated with local chains, and 63.9% are local independent hotel. This indicates a potential branding opportunity for local hotels in Sabah.

To compare the three categories in term of size, the number of rooms was used. The data reveals that 44.4% of the international chain hotels have between 101 and 200 rooms, where only 11% have less than 100 rooms. For the local chain, the majority of the Sabah hotels in this category have less than 100 rooms with a percentage of 73.3%, where only 3.3% of the hotels in this category have between 400 and 500 rooms. Finally, 78.3% of the independent hotels in Sabah have less than 100 rooms, but 2.9% have more than 500 rooms.

To further test the distribution of deviation among the three hotel categories, a normality test was performed. The results also show that both local chain $D(30) = .23, p < .001$, and independent hotels $D(69) = .23, p < .001$, are significantly non-normal. This indicates that there is no normality distribution within the two categories. The international chain hotels, $D(9) = .18, p > .05$, is found to be not significant, indicating that the distribution of the number of rooms in the international chain hotel is not significantly different from normal distribution. This result was expected because most international chain hotels are big in size and have larger number of rooms, especially when taking into account the small number of these hotels in Sabah with only 19.84% of the total number of rooms. Table 1 summarizes the results of this data.

TABLE 1: Descriptive Statistics for Sabah Hotels Number of Rooms

	Number of Hotels	Total Number of Rooms	M	Standard Deviation	K-S Test	df	Sig
International Chain	9 (8.3%)	1994 (19.84%)	221.56	129.56	.18	9	.20
Local Chain	30 (27.8%)	2870 (28.56%)	95.67	91.66	.23	30	.00
Independent Hotels	69 (63.9%)	5185 (51.59%)	75.14	105.65	.23	69	.00
Total	108 (100%)	10049 (100%)					

The results above give us the physical exposure of the hotels, which may not necessarily be the same in the virtual exposure in the Internet. The following sections discuss the virtual exposure of the hotel in the Internet.

Content Pages on the Website Sitemaps

Content is defined as the information features or services that are offered in the website (Zafiroopoulos & Vrana, 2006). The representation of information on the website can be measured either quantitatively or qualitatively (Law, Qi, & Buhalis, 2010). The qualitative measure indicates the nature of the content, whereas the quantitative measure indicates the amount of information provided. For a hotel website, studies suggest that the amount of information provided is more significant for travelers looking for information. Because in this paper we want to look into the quality of information, the number of content pages of hotel websites was investigated.

As shown in Table 2, the total web pages of Sabah hotel websites is 2720 web pages, where websites of the international chain hotels contains 500 web pages which represent 18.38% of the total number of web pages. Local chain hotels are represented by 718 web pages on their official websites, about 26% of the total number of the web pages. Finally, more than half of the total number of web pages represents the independent hotels with 1502 web pages. This shows that both independent hotels and local chain hotels are less represented on their own websites compared to the international chain hotels which are represented with well adequate number of web pages. This can be justified by two reasons. Firstly, the websites of the international

chain consists of unified web pages which display all information of affiliated hotels of the chain in various geographical areas. Secondly, most of the international chain hotels websites are presented with at least more than one language; this increases the frequency of the same information provided in multiple web pages on the same URL.

TABLE 2: Descriptive Statistics for Content Pages in the Websites

	Number of Hotels	Number of Content Pages	Mean	Standard Deviation	Variance	Skewness
International Chain	9 (8.3%)	500 (18.38%)	55.56	59.81	3577.27	2.17
Local Chain	30 (27.8%)	718 (26.39%)	23.93	27.82	774.20	1.86
Independent Hotels	69 (63.9%)	1502 (55.22%)	21.77	56.84	3185.68	7.04
Total	108 (100%)	2720 (100%)	25.19	50.86		

Web Pages Indexed by Search Engines

It is not only necessary to have a rich information website, but the more important is that this information can be reached through search engines. Search engines contain a robot called a spider which scans the keyword entered and crawls for any pages on the Internet contained that keyword. The well maintained web pages will appear on the first pages of the search engine. Although the crawling power of search engines differs depends on their spider algorithms, Google is recognised to be the strongest among other search engines. To assess how Sabah hotels’ websites are crawled by search engines robot, the number of web pages of each hotel’s website that appear in the search engine was recorded. The results of this process are presented in Table 3. From the results shown above, all 108 URLs of the hotels were indexed by the search engine. This is not surprising since the search engine will index any registered domain name. However, results revealed that the total number of web pages indexed by the search engine is 2904 web pages, where 20.38% belong to international chain hotels, 32.02% represent the local chain, and 47.54% of the total web pages represent the independent hotels. These results reveal that independent hotels are well indexed in the search engines compared to other two categories.

TABLE 3: Descriptive Statistics for Web Pages Indexed by Search Engine

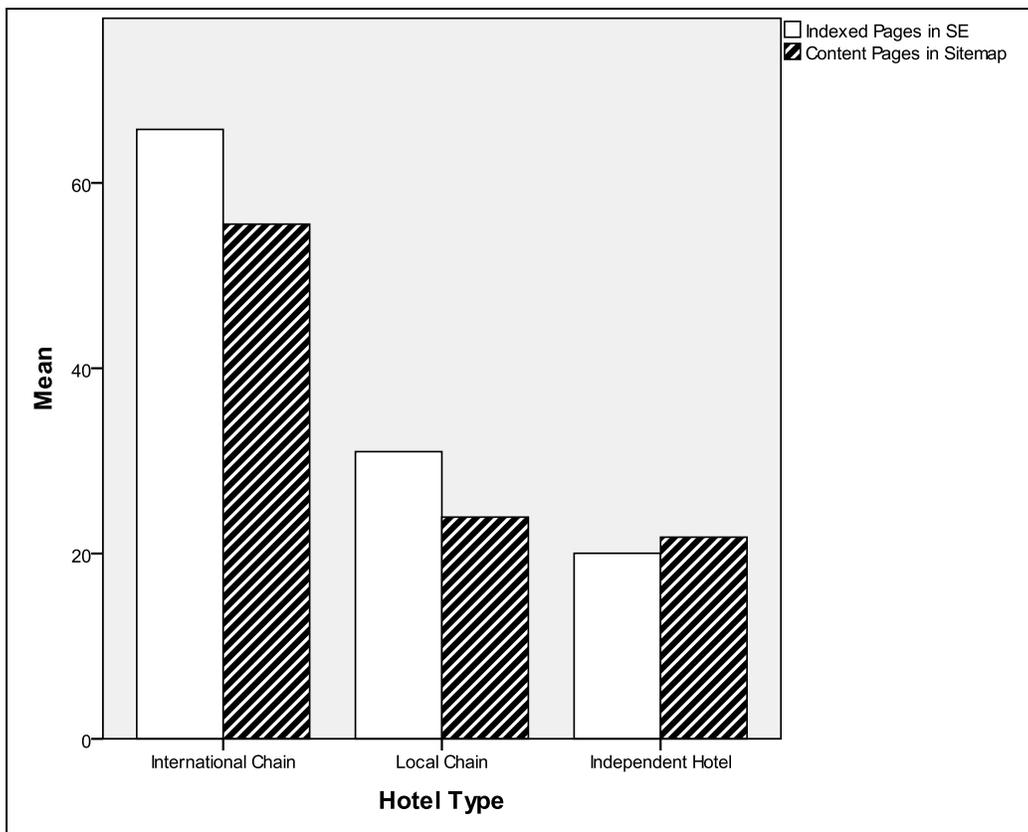
	Number of Hotels	Number of Indexed Pages	Mean	Standard Deviation	Variance	Skewness
International Chain	9 (8.3%)	592 (20.38%)	65.78	84.26	7099.94	1.59
Local Chain	30 (27.8%)	930 (32.02%)	31.00	38.81	1506.20	2.00

Independent Hotels	69 (63.9%)	1382 (47.58%)	20.03	56.14	3151.85	7.29
Total	108 (100%)	2904 (%100)	26.89	55.71		

Differences between Content Pages and Indexed Pages

For the ideal exposure of the web pages of a website on the Internet, it aimed to have the number of content pages equal to the number of pages crawled by a search engine. This mean that the website is well indexed, well updated and more importantly, the indexing update in the search engine is taking place quickly so no outdated content web page is still indexed after its removal from the website’s server. This will be clearly seen when calculating the index ratio. An initial look at the data as showed in Figure 1 illustrates that the number of content pages in the website is lower than the web pages indexed in search engine for both international chain and local chain hotels in Sabah, where it is the contrary for the independent hotels, where the number of the content pages in the sitemap is higher than the number of web pages indexed in the search engine.

FIGURE 1: Content Pages and Indexed Page in Search Engine



To further investigate whether an increase in number of content pages would lead to an increase of number of web pages indexed by search engine, a bivariate correlation was performed. From the results shown in Table 4, the number of content pages is positively

related to the number of web pages indexed in the search engine ($r = .93, p < .001$). This implies to the webmasters of the hotels website to increase the number of content pages in their websites if they want a good visibility of their web pages in the search engine. The search engine crawling for the hotels websites therefore accounts for 87% ($R^2 = .87$) of the variability in the number of web pages indexed in search engine.

TABLE 4: Correlation between Content Pages and Indexed Web Pages

		Content Pages	Indexed Web Pages
Contents Pages	Pearson Correlation	1	.935**
	Sig.		.000
	N	108	108
Indexed Web Pages	Pearson Correlation	.935**	1
	Sig.	.000	
	N	108	108
Mean		25.19	26.89
Standard Deviation		50.86	55.71

** . Correlation is significant at the 0.01 level (1-tailed).

Index Ratio

In order to better see the differences between the number of content pages on the hotel’s website and the number of web pages indexed by a search engine, the search engine index ratio was further examined using this formula:

$$Index\ Ratio = \frac{Total\ Number\ of\ pages\ indexed\ by\ a\ search\ engine}{Total\ Number\ of\ content\ pages\ on\ a\ hotel\ sitemap} \quad (1)$$

TABLE 5: Index Ratio for Sabah Hotel’s Website

Hotel Categories		Frequency	Percent	Valid Percent	Cumulative Percent
International Chain	Less than 1	5	55.6	55.6	55.6
	Equal to 1	4	44.4	44.4	100.0
	Total	9	100.0	100.0	
Local Chain	Less than 1	9	30.0	30.0	30.0
	Equal to 1	4	13.3	13.3	43.3
	Greater than 1	17	56.7	56.7	100.0
	Total	30	100.0	100.0	
Independent Hotels	Less than 1	35	50.7	50.7	50.7
	Equal to 1	11	15.9	15.9	66.7
	Greater to 1	23	33.3	33.3	100.0

Hotel Categories		Frequency	Percent	Valid Percent	Cumulative Percent
International Chain	Less than 1	5	55.6	55.6	55.6
	Equal to 1	4	44.4	44.4	100.0
	Total	9	100.0	100.0	
Local Chain	Less than 1	9	30.0	30.0	30.0
	Equal to 1	4	13.3	13.3	43.3
	Greater than 1	17	56.7	56.7	100.0
	Total	30	100.0	100.0	
Independent Hotels	Less than 1	35	50.7	50.7	50.7
	Equal to 1	11	15.9	15.9	66.7
	Greater to 1	23	33.3	33.3	100.0
	Total	69	100.0	100.0	

According to Leung & Law (2012), if the ratio is less than 1, it indicates that the hotels websites is not well indexed by the search engine, being the total number of pages indexed by a search engine is less than the total number of content pages on a hotel sitemap, This happen when the hotel's website are not making effort to increase their visibility on the search engines, by ignoring either keywords and metadata generation, or submitting the URL to the search engine directory. If however, the ratio is greater than 1, this means that the search engine have indexed more web pages than the actual pages existing in the website's sitemap. This is because the website may have been recently updated and the outdated web pages are still indexed in the search engine. As explained earlier, the indexing process in the search engine may take time to be put or to be removed. Another explanation for the greater index ratio is that some hotel's websites may still store older versions of their content pages on their servers, which they still can be crawled by the search engines. Figure 2 is an example of an existing outdated content page in the website.

FIGURE 2: An Example of an Outdated Indexed Content Page (June 2012)

Home Overview Guest Rooms Reservations Dining Meetings & Events Recreation **Promotion** Gallery Location Attractions Contact

Latest Promotions!

Value Stay
Valid till: 31/03/2011

Be it business or leisure, ~~XXXXX~~ Hotel in ~~XXXXXXXX~~ awaits you! Come stay with us for only RM150nett per peron.

[Book Now!](#)

Amadeus "We will rock you!"
Valid till: -

The live band is back and so is the FUN! Now we have theme night on every night of the week and the talk of the town is the "Beer Buffet" night on Friday where you can drink till you drop, literally...So come and join the FUN!

PLANTATION
ming
Lobby Lounge
AMADEUS
e-nice

As illustrated in Table 5, for the international chain hotels, the results show that 55.6% of their websites are less than 1, and 44.4% are equal to 1, which means that total number of pages indexed by the search engine is equal to the number of content pages on a hotel sitemap. No website in this category was found to have more content pages on their websites than those indexed by the search engine.

For the local chain hotels, the majority of their websites (56.7%) have an index ratio greater than 1, 30% of them are less than 1, and only 13.3% have an index ratio equal to 1. Finally, for the independent hotels, 50.7% of their websites have a ratio lower than 1, whereby 15.9% have an index ratio which is equal to 1, and 33.3% of their website have an index ratio greater than 1. What we can conclude from these results, is that international chain, have the majority of hotels which are well indexed, with a percentage of 44.4%. Although not a satisfying rate for international hotels, this may be because the websites' webmasters do not tend to regularly send their sitemaps to the search engines.

The advantage that these international hotels have over the other hotels is the crawling power of their URLs and domain names in the search engines due to the branding of their companies. This advantage has an impact on the fast update of web pages indexed in the search engine, where we do not see any invalid or outdated web pages, thus no index ratio greater than 1. 50% of the local chain index ratios are greater than 1. This means that there is a slow process of these web pages to be updated in the search engines, with many outdated and invalid web pages are still indexed and yet to be removed. This has a negative impact in the sense that it provides outdated information about these hotels. In this regard, Independent hotels are in a better position than the local chain hotels with only 33.3% of index ratios greater than 1. The independent hotels

are also in a better position than the local chain hotels in regards to index ratios equal to 1, where it is 15.9% for independent hotels compared to 13.3% for local chain hotels. This is relatively the ideal ratio where the number of both web pages in the sitemap is equal to the number of indexed pages in the search engine.

CONCLUSION

This paper examines the Sabah hotels' websites exposure on the Internet. The results revealed that for both international and local chain hotels, the number of content pages in the website is lower than the web pages indexed in the search engine. For independent hotels, the number of the content pages in the sitemap is higher than the number of web pages indexed in the search engine.

The results hold both advantages and disadvantages for the three categories. Webmasters of the independent hotels should consider the indexing performance of their websites, by sending a regular updates of their sitemaps to the search engines. While independent hotels lack a better presence in the crawling power of the search engine, international chain hotels fail to take this opportunity to illuminate the redundant web pages in their servers, which represent the largest margin among the three categories. Webmasters of these hotels should not only keep updating their websites, but they also need to remove outdated and invalid web pages from their servers so they will not be crawled by search engines. Independent hotels in the other hands are well indexed in the search engines but the lack of well maintained content web pages makes their websites less competitive and efficient as illustrated in Figure 5. This can be controlled by investing more on the quality of web contents in a more advanced design using the latest tools of web development.

One of the limitations of this paper is the structure designs of the hotels' websites. This comes in two perspectives. The first is the differences in the information arrangement between the websites, while the second is the multiple languages contents of some websites in which it makes the content pages in the website difficult to track. Nevertheless, this variation will not affect the overall process as websites of such contents are very minimal.

Further research may consider expanding this study to cover all hotels in Malaysian in order to know their exposure. Since this paper has only taken Google search engine, being the most powerful search engine and the most popular in Malaysia, further research may include more search engines to allow a comparison of results for a better maximization of visibility and popularity.

REFERENCES

- Buhalis, & Law. (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.
- Doolin, B., Burgess, L., & Cooper, J. (2002). Evaluating the use of the Web for tourism marketing: a case study from New Zealand. *Tourism Management*, (23), 557–561.
- Gretzel, U., Yuan, Y., & Fesenmaier, D. (2000). Preparing for the new economy: Advertising strategies and change in Destination Marketing Organizations. *Journal of Travel Research*, 39, 146-156.
- Hanna, J., & Millar, R. (1997). Promoting tourism on the Internet. *Tourism Management*, 18 (7), 469-470.
- Hashim, N. H., & Murphy, J. (2007). Branding on the web: Evolving domain name usage among Malaysian hotels. *Tourism Management*, 28 (2), 621–624.
- Hashim, N. H., Murphy, J., Purchase, S., & O'Connor, P. (2010). Website and email adoption by Malaysian hotels. *International Journal of Hospitality Management*, 29 (1), 194-196.
- Hernández, B., Jiménez, J., & Martín, M. J. (2009). Key website factors in e-business strategy. *International Journal of Information Management*, 29 (5), 362–371.
- Klein, B. (2002). When do users detect information quality problems on the World Wide Web? The 8th Americas Conference on Information Systems, (pp. 1101-1103).
- Knight, S.-a., & Burn, J. (2005). Developing a framework for assessing information quality on the World Wide Web. *Informing Science*, 8, 159-172.
- Law, R., Qi, S., & Buhalis, D. (2010). Progress in tourism management: A review of website evaluation in tourism research. *Tourism Management*, 31 (3), 297–313.
- Leung, R., & Law, R. (2012). Hotel information exposure in cyberspace: The case of Hong Kong. In M. Fuchs, F. Ricci, & L. Cantoni (Ed.), *Information and Communication Technologies* (pp. 132-142). Helsingborg, Sweden: Springer-Verlag\Wien.
- Lewandowski, D. (2008). A three-year study on the freshness of web search engine databases. *Journal of Information Science*, 34 (6), 817-831.
- Mullen, T. (2000). Travel's long journey to the web. *InternetWeek* (83), 103–106.
- Paraskevas, A., Katsogridakis, I., Law, R., & Buhalis, D. (2011). Search engine marketing: Transforming search engines to hotel distribution channels. *The Cornell Hotel Restaurant Administration Quarterly*, 52 (2), 200-208.
- Pew Internet & American Life Project. (2012). Search Engine Use 2012. Pew Research Center.
- Rouse, M. (2005). <http://searchsoa.techtarget.com/definition/site-map>.
- Zafiropoulos, C., & Vrana, V. (2006). A framework for the Evaluation of hotel website: The case of Greece. *Information Technology & Tourism*, 8, 239–254.