Marek Nowacki
WSB University of Poznan

World Cities' Image in TripAdvisor Users' Reviews

The purpose of the article is to identify five world cities' images by analysing the tourist attractions reviews posted on TripAdvisor. The phenomenological study of content analysis was used: first quantitative, and second qualitative. The analysis of the tourist attractions made it possible to identify their characteristic features and the attributes of the positive and negative attractions' image. Additionally, some specific attributes of the surveyed cities' images were identified. Study limitations include the limited reliability of the opinions published in TripAdvisor and a relatively small research sample.

Key words: content analysis, TextMining, TripAdvisor, destination image, tourist attractions

Marek Nowacki
The Faculty of Economics
ul. Powstańców Wielkopolskich 5,
61-895 Poznań
Phone: +49 61 655 33 33
Email: marek.nowacki@wsb.poznan.pl

Marek Nowacki is associate professor of tourism and recreation at the WSB University of Poznan with an interest in heritage and tourism. His current research focuses on role of social media in tourism as well as adventure tourism and recreation.

Introduction

By changing the distribution channels of information related to tourism the Internet has changed the way tourism is planned and consumed (Buhalis and Law, 2008). Nowadays the Internet is evolving from a medium that provides information to consumers into a medium where user-generated content (UGC) is the norm. The content generated by consumers for use by others has gained considerable credibility and has become an important and even main source of consumer information for tourism services (Jalilvand et al, 2012). Consumers trust much more UGC found on the Internet than other sources of information and believe that this content is created by honest, so-called ordinary consumers (other tourists) (Online reviews and endorsement, 2015). Currently, TripAdvisor is the most commonly used tool for evaluating and recommending tourism products with 456 million unique visitors per month. Around 200 new inputs, i.e., reviews, comments and photos are being uploaded every minute in relation to 7.7 million hotels, restaurants and tourist attractions in 49 countries around the world (About TripAdvisor, 2018).

Destination image is a type of attitude that consists of individual beliefs, feelings and general impressions of a destination (Baloglu and McCleary, 1999). Some researchers define the image as a set of attributes (Gartner 1989), others do so holistically (Um and Crompton, 1990). Tourist attractions are a functional attribute of the tourism destination image (Gallarza, Gil and Calderón, 2002; Echtner and Ritchie, 2003). A review of the image studies reveals that it has three dimensions: functional-psychological, attributes-holistic and common-unique (Echtner and Ritchie; 2003). Many researchers also distinguish three (or four) components of the image (Gartner, 1993; Baloglu, 1999): cognitive (including knowledge and beliefs about destination attributes), affective (emotional sensations triggered by destination), conative (behavioural activities related to the destination, like visit or recommendation) and holistic (general evaluation of the destination).

Up to date several authors have been analysing images of destinations in TripAdvisor and other social media. For example, Kladou and Mavragani (2015) have analysed Istanbul's image on the basis of TripAdvisor posts on the "Historic Areas of Istanbul" page. They've assessed the cognitive, affective and conative components of the image. In another study Tamajon and Valiente (2015) have compared the images of Barcelona created by users on the TripAdvisor Forum and ones created on traditional DMO websites. The authors identified factors, typologies and components specific for DMO websites and other social media platforms. Marine-Roig and Clave (2015) have analysed the image of Catalonia using user-generated content. More than 130,000 trip diaries (travel blogs and online travel reviews) have been analysed. Buonincontri, Cinquegrani and Martone (2017) have analysed online reviews posted on TripAdvisor on attractions in Naples (Italy). Toral, Martinez-Torres and Gonzalez-Rodriguez (2017) have identified unique attributes of tourist destination using text mining and analysis of variance. Data were collected form eWOM community Ciao UK. Smith at al. (2018) identified the image of the Balkan region, by content analysis of comments posted on TripAdvisor. They identified key elements of the Balkan region's image, its emotional elements, role of locals and the complexity of tourism experience.

There is an existing research gap between identification of activities undertaken by tourists visiting the destination and a more comprehensive view of the image of the destination through the analysis of opinions on tourist attractions. Hence, the aim of the study is to compare selected attributes of destination images of world cities, by analysing the comments posted by TripAdvisor users which are related to tourist attractions.

Methodology

This study used content analysis of comments posted on TripAdvisor.co.uk. In the first phase the STATISTICA package TextMining analysis tool was used. The words occurring in

the comments most frequently were calculated. In the next step Singular Value Decomposition (SVD) transformation was performed, and the relationship between the words was illustrated on the graphics. Comments referred to 50 different tourist attractions from five cities: Tokyo, New York, London, Dubai and Warsaw were analysed. The selection of cities and attractions was purposeful: the largest tourist cities in Asia, America and Europe (Poland – the author's country), and the highest ranked attractions in these cities. For every attraction, the sample group of 50 positive opinions (attractions rated as "excellent") and 50 negative opinions (attractions rated as "terrible") were downloaded. However, in many cases such sampling was not possible because the attractions had not received so many negative comments. In these situations less extreme opinions ("poor") were also downloaded. However, in 23 cases it was not even possible to obtain 50 "terrible" or "poor" reviews. Since the purpose of the research was mainly qualitative analysis rather than quantitative, it was decided to leave the asymmetric sample (with a predominance of positive opinions). Reviews were downloaded using Web Scraper (webscraper.io) between May 2nd and 12th 2017. A total of 4742 reviews were received.

Data analysis, results and discussion

The analysis of the results began by counting the words used most frequently in all reviews. The words most common are: *visit* (1379), *see* (1061), *museum* (946), *place* (929), *time* (913) and *park* (804). These are attributes of the destination image's cognitive component and their content reflects the structure of the attractions that were included in the sample: most of them were museums (19) and parks (9). The main activities that visitors describe are *visit*, *see* and *walk*, which are the attributes of the behavioural dimension of the destination image. There are also many attributes of the affective component in the opinions, these are: *great, beauty, love, amaze, good, nice, enjoy* and *interest*.
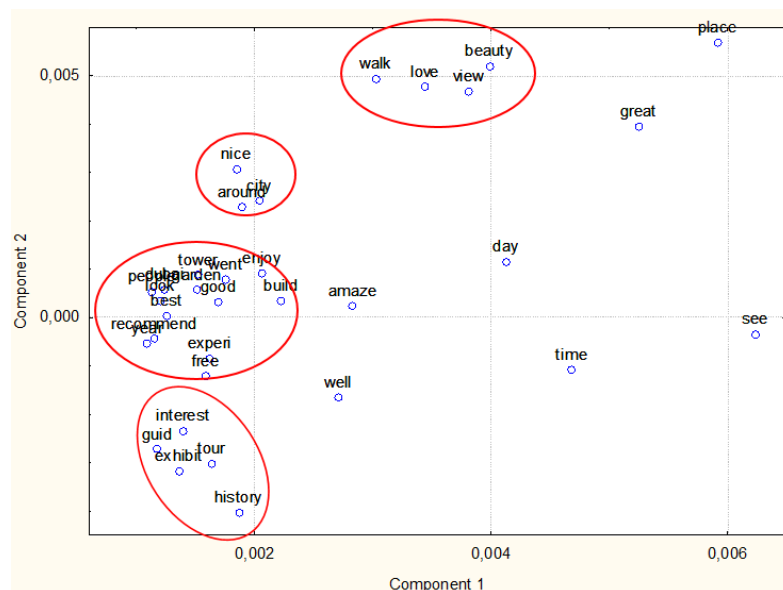
Figure 1. Scatter plot of the values of the first two components for words related to positive opinions

In order to find the connections between the analysed words, to determine the semantic structure of opinions and to reduce the number of variables describing the analysed set of words, SVD (Singular Value Decomposition) transformation was performed (Demski, 2006). The purpose of this technique is to reduce the overall dimensionality of the input matrix to a lower-dimensional space, where each consecutive dimension represents the possible largest degree of variability (between words) (StatSoft, Inc., 2007). The relationships between the words were examined by the graphical analysis of word value coefficients plotted using the first two components (Demski, 2006). Figure 1 shows that the analysed words formed several distinct clusters. The first cluster is made up of such words as *history, exhibit, guide, tour* and *interest*, which are attributes of the cognitive dimension of the image. The second cluster that is made up of the words *walk, love, view, beauty* refers to walks accompanied by admiration of beauty (cognitive-affective attributes). The third cluster that is made up of the words *tower, garden, build, went, good, enjoy, look, best, people* refers to satisfaction connected with the experience of visiting gardens, buildings and view towers, worth recommending and often made available to visitors for free (cognitive-affective attributes). The last and least numerous

cluster that is made up of the words *nice, around, city* concerns people who are satisfied with the city tour.

Particularly interesting for image analysis is the identification of attributes that illustrate the affective and coactive components, as these two dimensions most affect the intension to travel to the destination (Wang and Hsu, 2010). The largest number of positive attributes can be found in opinions relating to London tourist attractions (n = 866), which means more than 1.5 words per one review ($\bar{x}$ = 1.527) (see Table 1). Relatively the fewest of the positive image attributes can be found in opinions about Tokyo attractions: the word beauty appears as many as 125 times. Relatively the lowest number of words of the positive image attributes can be found in the opinions about the Dubai attractions (n = 659, $\bar{x}$ = 1.177) and New York (n = 661, $\bar{x}$ = 1.200).

Table 1: The most common words-attributes of the positive destination images

Word	Tokyo	New York	London	Warsaw	Dubai
great	123	85	134	113	109
beauty	125	106	63	121	57
love	63	92	126	58	54
amaze	51	95	81	40	96
well	63	58	89	69	63
enjoy	61	38	59	49	55
nice	66	35	32	73	43
good	49	32	55	41	63
free	64	29	64	27	20
interesting	40	0	74	59	0
best	22	34	32	26	52
recommend	36	27	33	34	23
wonder	24	30	24	27	24
Sum	787	661	866	737	659
Positive / all ($\bar{x}$)	1.398	1.200	1.527	1.328	1.177

In the case of negative opinions, SVD transformation also was performed and the relationships between the words were graphically presented in the scatter plot using the first two components (Figure 2). The first cluster is made up of the words *go, queue, line wait, day, view, people* which are connected to queuing for attractions, waiting to enter and with people (cognitive-affective dimension of the image). The second cluster of words is connected with dissatisfaction associated with high entry prices (*expense, money*), low value and no interest in the attraction exposition (*nothing, nice, interest*), as well as crowdedness and closed access (*crowd, minute, close*). The third cluster is composed of the words: *disappoint, look, walk, tower, staff, build*.
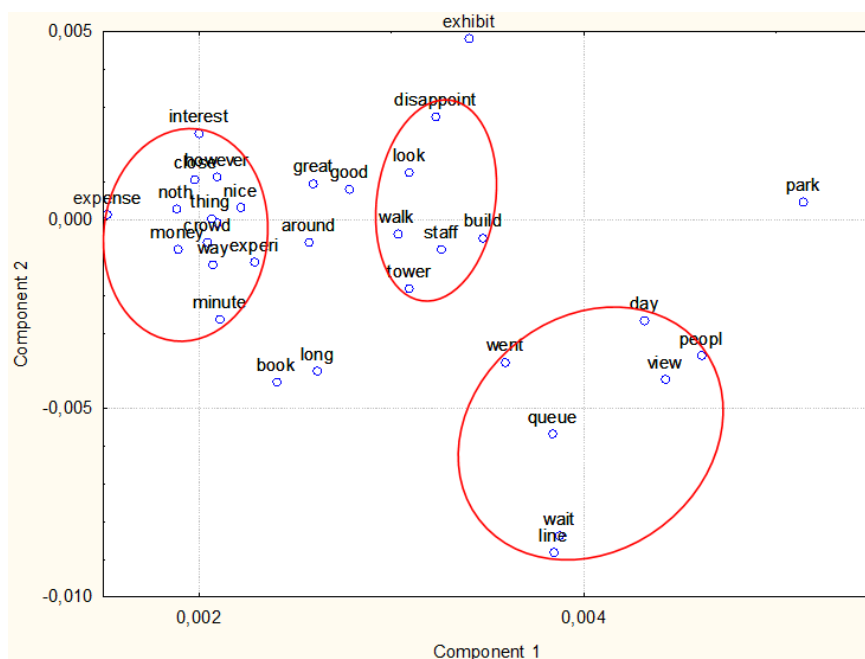

Figure 2. Scatter plot of the values of the first two components for words related to negative opinions

The largest volume of words with negative connotations can be found in the opinions about New York's tourist attractions (n = 505, $\bar{x}$ = 0.973) (see Table 2). The most common word is *line* (105), which indicates that the biggest problem of the image of New York is waiting in queues. There are no such problems in Warsaw, for example, where the words *crowd, line, waste* are absent from opinions. However, there are referrals to queues as many as

24 times. Warsaw's attractions are also distinguished by the low price of admission – there are no such words as *expensive, money and price*. In turn, the high prices of attractions distinguish Dubai in a negative way (*expensive = 36, money = 40, price = 36*). It is worth pointing out that while the opinions concerning Tokyo's tourist attractions have the highest density in terms of negative words (average per comment $\overline{x} = 1.07$), the number of negative comments is relatively small (259) compared to other cities. The only word that is found in the opinions about Tokyo's attractions much more often than in others is *crowd*, which is understandable in the world's largest metropolis with over 34 million inhabitants.

Table 2: The most common words-attributes of the negative destination image

Word	Tokyo	New York	London	Warsaw	Dubai
wait	25	61	27	14	52
disappoint	22	43	43	31	38
queue	26	48	46	24	25
line	22	105	17	0	23
long	19	44	23	12	23
close	20	24	29	20	24
money (not worth)	16	29	32	0	40
crowd	35	19	28	0	30
interest (not)	11	18	50	31	0
expensive	16	22	30	0	36
price (high)	0	26	30	0	36
waste (time)	11	21	24	0	30
poor	0	17	19	12	21
Sum	259	505	419	163	659
Negative/all ($\overline{x}$)	1.070	0.973	0.848	0.685	1.177

The analysis has shown that comments posted on TripAdvisor concerning tourist attractions can be used to identify the attributes of a destination image: all three cognitive (e.g. *museums, sites, urban parks, sightseeing, viewing and walking*), affective (e.g. *great, beautiful*) and conative components (e.g. *recommend*). The most frequently used words in the opinions

of TripAdvisor users were related to cognitive component attributes, which was also found in studies looking at other social media (Xiang and Gretzel, 2010).

The analysis of the links between the word-attributes, broken down into positive and negative comments, has allowed to reveal the context of the situations that causes such opinions. Situations that are conducive to the formation of positive opinions include: visiting historical expositions (guided tours), accompanied walks with admiring views or visiting gardens, buildings and viewpoints. Negative situations include, for example, waiting in queues for visiting attractions or dissatisfaction caused by high entrance price or a low value of attractions. However, it is not a single experience, but their sum that determines the image of the destination (and especially its affective component): the image is a result of positive and negative attitudes that an individual has towards particular products and services (attractions, hotels, restaurants, tours) experienced during the stay in a destination (Leisen, 2001). This confirms the necessity of cooperation and coordination of all tourism development stakeholders (managers of visitor attractions, restaurants, hotels) in the field of image shaping and destination branding (Konecnik and Gartner, 2007).

Limitations of this paper include a relatively small sample and limited credibility as opinions are only taken from a portal dedicated to users from United Kingdom (TripAdvisor.co.uk).

References

------ (2018). "About TripAdvisor". Available at <https://tripadvisor.mediaroom.com/pl-about-us> (10 May 2018)

Baloglu, S. & McCleary, K.W. (1999). A model of destination image formation. *Annals of Tourism Research*, 26, 868–897.

Baloglu, S. (1999). A path analytic model of visitation intention involving information sources, socio-psychological motivations, and destination image. *Journal of Travel and Tourism Marketing*, 8, 81–91.

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, 609–623.

Demski, T. (2006). Czy leciał z nami pilot, czyli text mining na przykładzie opisów wypadków lotniczych. Available at <http://www.statsoft.pl/Portals/0/Downloads/Text_mining_wypadki_lotnicze.pdf> (15 May 2018)

Echtner, C.M. & Richie, B.J.R. (2003). The Meaning and Measurement of Destination Image. *The Journal of Tourism Studies*, 14, 37–48.

Gallarza, M., Gil, I. & Calderón, H. (2002). Destination image. Towards a conceptual framework. *Annals of Tourism Research*, 29, 56–78.

Gartner, W.C. (1993). Image formation process. *Journal of Travel and Tourism Marketing*, 2, 191–215.

Jalilvand, M.R., Samiei, N., Dini, B. & Manzari, P.Y. (2012). Examining the structural relationships of electronic word of mouth, destination image, tourist attitude toward destination and travel intention: An integrated approach. Journal of Destination Marketing and Management, 1, 134–143.

Kladou, S. & Mavragani, E. (2015). Assessing destination image: An online marketing approach and the case of TripAdvisor. *Journal of Destination Marketing & Management*, 4, 187–193.

Konecnik, M., & Gartner, W. (2007). Customer-based brand equity for a destination. Annals of Tourism Research, 34, 400–421.

Leisen, B. (2001). Image segmentation: the case of a tourism destination. *Journal of Services Marketing*, 15, 49–66.

Lluís Garay Tamajón & Gemma Cànoves Valiente (2017) Barcelona seen through the eyes of TripAdvisor: actors, typologies and components of destination image in social media platforms, *Current Issues in Tourism*, 20, 33-37,

Marine-Roig, E., & Anton Clavé, S. (2016). A detailed method for destination image analysis using user-generated content. *Information Technology & Tourism*, 15, 341–364.

------ (2015). "Online reviews and endorsement". Report on the CMA's call for information. Available at <https://www.gov.uk/cma-cases/online-reviews-and-endorsements> (10 May 2018)

Simeon, M. I., Buonincontri, P., Cinquegrani, F., & Martone, A. (2017). Exploring tourists' cultural experiences in Naples through online reviews. *Journal of Hospitality and Tourism Technology*, 8, 220–238.

Smith, M., Sulyok, J., Jancsik, A., Puczkó, L., Kiss, K., Sziva, I., Papp-Vary, A.F. & Michalkó, G. (2018). Nomen est omen – Tourist image of the Balkans. *Hungarian Geographical Bulletin*, 67, 173–188.

------ (2007). "STATISTICA (data analysis software system), version 8.0". StatSoft, Inc., Available at <www.statsoft.com> (15 May 2018)

Toral, S. L., Martínez-Torres, M. R., & Gonzalez-Rodriguez, M. R. (2018). Identification of the Unique Attributes of Tourist Destinations from Online Reviews. *Journal of Travel Research*, 57, 908–919.

Um, S. & Crompton, J.L. (1990). Attitude determinants in tourism destination choice. *Annals of Tourism Research*, 17, 432–448.

Wang, C. & Hsu, M.K. (2010). The Relationships of Destination Image, Satisfaction, and Behavioral Intentions: An Integrated Model. *Journal of Travel & Tourism Marketing*, 27, 829–843.

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, 31, 179–18