WHO ARE THE TOURISTS SHARING CONTENT ON SOCIAL MEDIA? BEHAVIOUR AND CHARACTERISTICS

Nuno M.B. Antunes[1]
Célia M.Q. Ramos[2]
Carlos M.R. Sousa[3]

ABSTRACT

Increasingly, tourist information on the Internet comes from the content generated by users, being no longer limited to the communications of tourism professionals and government entities. At the same time, the comments, opinions and reviews made by consumers are the type of information which tourists value the most. This reality means that the Internet has become an extremely valuable source of information for prospective travelers. This study aims to identify and characterize the behavior patterns of tourists on social media during and after their travels, specifically: who are the tourists who create content, what leads them to have this initiative, and what behaviors stand out on each of these platforms. The methodology is based on a questionnaire survey, which was applied to tourists. The data was analyzed using descriptive statistical methods, cluster analysis and principal component analysis. It is possible to categorize tourists according to three perspectives regarding use of social media during and after the trip: those who prefer review websites, those who prefer social networks and those who don't particularly use social media in general.

Keywords: Internet, Social Media, Social Networks, Review Websites, User-Generated Content, Tourist-Generated Content.

JEL Classification: L83, L86, M31, O33, Z33

1. INTRODUCTION

Technological advances have affected human life in a number of ways, one of which regards how information is accessed. The strong development of information and communication technologies (ICT) and the greater ease of access to the Internet have brought new means of interaction and coexistence to society (Damian, 2014), expressed through platforms such as social networks, personal blogs, (such as YouTube and Flickr), online communities, and others (Xiang & Gretzel, 2010). This reality forced marketers to consider a new form of communication: Digital Marketing, focused on directing marketing efforts towards connecting with the vast number of consumers who are constantly present online (Damian, 2014; Ryan, 2014).

At the same time, it is a fact that the evolution of ICT directly influences the tourism industry, given the nature of the tourism product itself, which, among other characteristics, is intangible and requires that the purchase be made in a significant period of time prior to consumption (Buhalis & Law, 2008). Because of these inherent traits, consumers often seek a better understanding of the tourism product during the decision-making stage, which

[1] Masters in Tourism Marketing, University of Algarve, Faro, Portugal (nunantunes@gmail.com)
[2] Research Centre for Spatial and Organizational Dynamics. ESGHT, University of Algarve, Faro, Portugal (cmramos@ualg.pt)
[3] ESGHT, University of Algarve, Faro, Portugal (cmsousa@ualg.pt)

highlights the great importance of information in this industry. Knowing this, companies and institutions in the area tend to favor online communication as a way to broadcast the quality of their tourism product, thus encouraging its consumption.

Pan, MacLaurin and Crotts, (2007), Carrera, Chiu, Pratipwattanawong, Chienwattanasuk, Ahmad and Murphy (2008), Litvin, Goldsmith and Pan, (2008), Yoo and Gretzel, (2008a), Yoo and Gretzel, (2008b), Yoo and Gretzel, (2008c), Gretzel et al., (2009), Yoo, Lee and Gretzel, (2009), Yoo, Lee, Gretzel and Fesenmaier, (2009), Xiang and Gretzel, (2010), Fotis, (2015) and Yoo and Gretzel, (2016) find that a growing share of tourism content available on the Internet comes from tourists themselves. At the same time, prospective travelers are increasingly basing their travel plans on content generated by other consumers (Buhalis & Law, 2008; O'Connor, 2008; Adjei, Noble, & Noble, 2010; Xiang & Gretzel, 2010, Munar & Jacobsen, 2014; Fotis, 2015).

In this context, social media — an extensive source of tourism content generated by consumers — are the focus of this study, which intends to provide researchers, tourism professionals and tourists alike with an in-depth analysis of the behavior of the tourist as a creator of content on the Internet.

Taking the purpose of the study into account, this article is structured in three sections, in addition to the present introduction and the conclusion. The first section, encompassing the literature review, aims to identify, describe and adequately substantiate the main issues surrounding the study of tourist behavior on social media. The methodology section describes the procedures adopted during the measurement of results based on the empirical data which was obtained via survey. The third section covers the analysis performed on the empirical data, including the most conclusive statistically verified results.

2. LITERATURE REVIEW

The evolution of ICTs has contributed to the emergence of social media, a type of platform for social interaction which can be characterized by its ease of use, the democratization of tools for sharing different types of content (multimedia, text, hyperlinks, among others), and the very nature of this medium, which enhances communication between users and minimizes physical barriers by facilitating interactions between different spaces and times.

The concept of social media, according to Munar and Jacobsen (2013: 3-4), is a term that covers various types of online platforms, among which, to name some of the most popular ones, are wikis (Wikitravel), blogs (Travelblog), microblogs (Twitter), social networks (Facebook), multimedia sharing websites (YouTube, Flickr, Instagram) and review websites (TripAdvisor). On the other hand, Fotis (2015) argues that social media can be defined as a term covering all online applications whose main function is the development and exchange of user generated content (UGC).

In the present study, and considering the goals of the investigation, websites where UGC exchange serves a secondary purpose were also considered; in other words, not only are websites like TripAdvisor considered, but also platforms in which reviewing is a secondary feature (as is the case of Booking.com). This decision was based on the reality that tourists consult all kinds of platforms where users generate content, regardless of whether this is a primary feature or not. This consideration allows for the analysis of a more comprehensive set of platforms where tourists can create or consume tourist information, which is the main objective of this study.

Social media have rapidly evolved over a short period of time, not only in functionalities, but also in users. Facebook, founded in early 2004, has an average of about 1.8 billion unique active users per month (Facebook, 2016). Twitter counts about 313 million users

(Twitter, 2016), TripAdvisor nearly 390 million (TripAdvisor, 2016) and Instagram over 500 million unique active users per month (Instagram, 2016).

The potential of social media has contributed to the emergence of new channels of communication between consumers and businesses. As far as tourism is concerned, social media are used as a platform for sharing opinions and experiences, through the creation of testimonies, which can be shared and consulted with other tourists.

In addition, tourism information content is increasingly available on the internet, a great deal of which is produced by tourists themselves, and represents an extremely valuable source of information for both consumers and tourism professionals (Pan, MacLaurin & Crotts, 2007; Carvalho et al., 2008; Litvin, Goldsmith & Pan, 2008; Yoo & Gretzel, 2008a; Yoo & Gretzel, 2008b; Yoo & Gretzel, 2008c; Gretzel et al., 2009; Yoo, Lee, Gretzel & Fesenmaier, 2009; Xiang & Gretzel, 2010; Fotis, 2015; Yoo & Gretzel, 2016).

As an essential element to allow frequent access to the Internet and social media while traveling, is the smartphone, which has already become one of the most important devices for tourists (Charlesworth, 2009), either to consult information, share experiences, or simply to keep in touch with the rest of the world (Kang & Schuett, 2013; Fotis, 2015; Yoo & Gretzel, 2016). In fact, a study by Tnooz (2010) revealed that technological solutions that allow access to social media are the most used during the tourist trip (38% in the US and 64% outside the US).

In the words of Yoo & Gretzel (2016: 192), "It is obvious that social media provides a fertile place for travelers to create and share their travel experiences and also take on an important role in tourist information and decision-making." It is increasingly relevant to understand the behavior of the tourist in relation to social media, as well as the motivations and factors that contribute to the interaction with these platforms during the tourist trip, regarding the sharing of content.

The main factors that motivate tourists to express themselves through social networks (Fotis, 2015) are: convenience, that is, the fact that sharing content on social media is an easy and accessible way to communicate while also storing content (experiences) online; showing off, in the sense that the tourist enjoys the idea that their acquaintances want to do or know what he/she is doing; sharing experiences and stories that truly deserve to be told; inviting others to live the experience, by recommending a certain destination or activity to their acquaintances; sharing happy moments with friends and acquaintances; expressing sociability, as nowadays it is expected of tourists to share some type of travel-related content while traveling, and tourists wish to act in conformity with this societal norm.

2.1 Tourists, participants and creators of social media content

Depending on the motivations that lead to the use of social media, a tourist — in this case, a user — shows different behaviors which characterize him depending on how he interacts with online content.

Tedjamulia, Dean, Olsen, and Albrecht (2005), Shao (2009), and Van Dijck (2009) indicate that there are three ways of interacting with content on social media: as a viewer, who only consumes content created by others; as a participant, who initiates consumer-to-consumer and consumer-to-content interactions, including comments, content reviews, and so on; and as content creator, which includes the effective creation of online content, whether in text, image, audio or video format.

The vast majority of social media users are just viewers, watching and interpreting all kinds of content produced by third parties on the many different platforms available online. As they gradually become involved in their respective online communities, some users evolve

from viewers to participants, and only a few of these participants eventually contribute with user-generated content (UGC).

Yoo and Gretzel (2016), supported by the existing literature (Acar & Polonsky, 2007; Lenhart, Madden, Macgill & Smith, 2007; Yoo, Lee & Gretzel, 2007; Jones & Fox, 2009; Nielsenwire, 2009; Yoo et al. 2009; Burgess, Sellitto, Cox & Buultjens, 2011), consider that the social interaction rate of social media users is strongly influenced by factors such as motivations in life, preferred types of social media, preferred types of device, level of trust in the community, user personality, among others. In addition to these factors, certain sociodemographic characteristics also impact use of social media: younger users tend to participate more actively and create more content; adult men contribute with content more often than adult women; among pre-adolescents, adolescents and university students, women create more content than men within the same age group.

2.2 Characterization of online tourists

In terms of the use of social media for tourism travel, several studies (Yoo, Lee & Gretzel, 2007; Yoo & Gretzel, 2008a; Yoo & Gretzel, 2008b) suggest a correlation between certain sociodemographic characteristics and certain behaviors on social media, In particular, those who produce social media content while traveling tend to be young, male, high-income, internet-savvy adults who travel frequently and are heavily involved in travel planning; in addition, users who share photographs of their tourism experiences tend to be of younger generations, while *baby boomers* (people born between 1946 and 1964) and the elderly typically to avoid sharing photos on the Internet.

In a more recent study, Yoo and Gretzel (2016) find that only about 20% of online tourists contribute with UGC. This means that the overwhelming majority of individuals only consume tourism UGC, most of which is created by only a small portion of all online tourists. Because of this, it is pertinent not only to assess who uses social media for purposes related to tourist travel, but also to ascertain who are the online tourists who, in fact, contribute with UGC on social media.

Through their study, Yoo and Gretzel (2016) verified that gender does not influence the propensity to use social media during the trip. In contrast, the age group of the users shows a great influence towards this propensity, as younger tourists use social media more often during the trip.

The online tourists who create UGC have a similar profile, as younger tourists are more likely to share content related to the tourist experience, compared to the older ones. It is estimated that this is the consequence of a lower involvement of the upper age groups in online communities, due to a lack of trust (Yoo et al., 2007; Burgess et al., 2011) and lesser time spent interacting with those communities. At the same time, the same influence of age is seen in the propensity to share photographs of the tourist trip: these mainly represent the content created by younger users, while older ones prefer text content. Tourists who are working full-time are also more likely to contribute with UGC than those in other occupational situations: this can be explained by the fact that full-time workers tend to live under a routine, so when faced with experiences that contradict this routine, they are more likely to feel the need to register and share those moments which characterize their escape from everyday life.

3. METODOLOGY

In order to find answers to the questions in this study, seeking to fulfill the initially stipulated research objectives, and also as a way to contribute to the scientific bibliography on the use

of social media in the tourist trip, the collection of first-hand data was deemed necessary. Thus, a quantitative approach was chosen and a survey was carried out on the population residing in any European Union country that has undertaken at least one tourist trip in the years 2014 and/or 2015.

It should be noted that, initially, the research universe was intended to include tourists residing in any European country, in order to ascertain conclusions about the behavior of European tourists. However, after analyzing the collected data, it was verified that 92% of the respondents reside in Portugal, so a decision was made to switch the angle of the analysis, focusing on a more detailed understanding of the universe of Portuguese tourists, rather than a broad description of the vast universe of European tourists, thus achieving greater trustworthiness in regards to the results of the study (see Table 1).

Table 1. Technical data sample

Research Universe	Portuguese tourists who traveled in the years 2014 or 2015
Sample size	250 valid questionnaires
Method of obtaining data	Survey via online questionnaire
Sample procedure	Sample by convenience
Start of the survey	June 20, 2016
End of survey	July 23, 2016

Source: Own Elaboration

3.1 Survey

The questionnaire was built around three main topics: the characterization of the respondent's tourist profile, the study of their social media consumption habits and perceptions regarding social media, and the demographic characterization of the respondent.

3.2 Sample and procedures

With the original research universe in mind, which, as mentioned before, included all European citizens who have traveled at least once in the years 2014 and/or 2015, the survey was carried out in such a way as to allow the participation of individuals located in any country.

As such, and also taking the resources available for this research into account, personally collecting answers would not be an efficient solution.

Thus, the survey was applied on the Internet, and the questionnaire was disseminated online, mainly through Facebook, and the sample was collected by convenience. This route of dissemination may represent a limitation, since it is implied that the respondents who gained knowledge of the survey via Facebook will have a favorable interaction with this platform and, as such, may have, *a priori*, a positive predisposition towards the use of social networks and/or the Internet in general.

The survey was conducted in three languages (Portuguese, English and French) and ran between June 20, 2016 and July 23, 2016, with a total of 335 questionnaires collected. The data was then organized and analyzed using IBM SPSS Statistics 20.

Of the 335 respondents, 64 were excluded because they did not travel in 2014 or 2015 and/or because they were not residents in the European Union. Of the 271 remaining respondents, 21 were excluded because they did not reside in Portugal. Concluding, the sample under study comprises 250 Portuguese individuals who have made at least one tourist trip in 2014 or 2015.

After the construction of the database, a statistical study was carried out in order to characterize social media consumption habits, through descriptive statistical analyzes and, in certain cases, cross tabulations. In addition, multivariate analyzes were carried out, namely Principal Component Analysis (PCA), so as to further study the question regarding the use of social networks versus review websites as tools to support decision making during the trip, and also Cluster Analysis, so as to segment the study sample according to its behavior on social media during and after the trip, followed by the characterization of each of the segments.

3.3 Demographic characterization of the sample

Table 2 shows the demographic characteristics of the sample according to the data obtained from the survey.

Table 2. Demographic characterization of the sample

Gender	
Male	32%
Female	68%
Age	
15 to 25 years	28%
26 to 35 years	11%
36 to 45 years	26%
46 to 55 years	23%
56 to 71 years	12%
Minimum age observed	15 years
Average of sample ages	39 years
Maximum age observed	71 years
Professional situation	
Unemployed	3%
Employee on behalf of others	61%
Self-employed	10%
Student	22%
Retired	4%
Number of years of schooling successfully completed	
Up to 9 years (basic education)	4%
Up to 12 years (secondary education)	16%
Up to 15 years (licentiate post-Bologna)	18%
Up to 17 years (pre-Bologna degree / masters degree)	38%
18 or more years (doctorates, post-graduates, second degrees, etc.)	24%
Number of household elements.	
1	16%
2	18%
3	33%
4	28%

5	4%
6 or more	1%
Average monthly net income of the household	
Less than € 1500	31%
Between 1500 € and 3000 €	49%
Between 3001 € and 4500 €	16%
More than € 4500	4%

Source: Own Elaboration

About 68% of respondents are female, which may be a consequence of the strong presence of the survey on Facebook, where some women respondents seem to have contributed significantly to the dissemination of the questionnaire. The sample is aged between the 15 and 71 years old, with the average being 39 years old, 38% of the sample completed 17 years of education (which corresponds to the current master's degree in Portugal), 33% belong to a household with 3 elements and 49% of the sample has an average monthly net income of 1,500 € to 3,000 € in the respective household.

4. ANALYSIS OF EMPIRICAL DATA

4.1 Characterization of tourists who create online content

The answers given by the sample indicate that about 73% of tourists create content online, either in social networks or in review websites.

Based on the works of Yoo, Lee and Gretzel (2007), Yoo and Gretzel (2008a) and Yoo and Gretzel (2008b), it was possible to piece together a general characterization of tourists contributing with UGC on social media (see Table 3), as well as a characterization of those who share their travel pictures online (see Table 4).

Table 3. Characterization of tourists creating online content

Characteristics	Summary
Gender	Male
Age	Young adults
Yield	High
Internet domain	High
Travel frequency	High

Source: Own elaboration, based on Yoo, Lee and Gretzel (2007), Yoo and Gretzel (2008a) and Yoo and Gretzel (2008b).

Regarding the sociodemographic characterization of tourists that generate content on social media, the results obtained in the present study generally correspond to those indicated by Yoo, Lee and Gretzel (2007), Yoo and Gretzel (2008a) and Yoo and Gretzel (2008b); however, there are some discrepancies, namely:

1. Gender seems to have little or no influence on the propensity to generate content, contesting the hypothesis that men generate most of the content;
2. Although a large proportion of respondents between 15 and 25 years of age (93%) generate content on social media, this behavior also occurs (in a slightly larger number)

in the 26-35 age group (96%) and from 36 to 45 years old (97%), challenging the hypothesis that young adults are the ones who generate most of the content;

3. Respondents with monthly net household incomes between € 3,001 and € 4,500 show a greater tendency (97%) towards sharing content on social media than respondents with a household income greater than € 4,500 (77%). However, it should be noted that respondents with an income above € 4,500 are mainly people over 55, which would explain the lower predisposition towards social media sharing.

Table 4. Characterization of tourists creating content in social media

Characteristics	Share tourism content in social media (%)
Gender	
Male	92%
Female	91%
Age	
15 to 25 years	93%
26 to 35 years	96%
36 to 45 years	97%
46 to 55 years	89%
56 to 71 years	85%
Average monthly net income of the household	
Less than € 1500	94%
Between 1500 € and 3000 €	91%
Between 3001 € and 4500 €	97%
More than € 4500	77%
Familiarity with social media	
1 (not at all familiar)	90%
2	83%
3	93%
4	92%
5 (totally familiar)	100%
Number of travels made in 2014/2015	
1	84%
2	89%
3	96%
4	94%
5	96%
Between 6 and 10	94%
11 or more	100%

Source: Own Elaboration

As for the concordances, the results show that a greater degree of familiarity with social media leads to a greater propensity to generate content, confirming the hypothesis that a greater dominance of the Internet is a positive factor towards content sharing.

At the same time, people who travel more often also show a greater tendency to share content in social media, which also validates the results presented by Yoo, Lee and Gretzel (2007), Yoo and Gretzel (2008a) and Yoo and Gretzel (2008b).

4.2 Principal component analysis

As part of the data analysis, a more thorough study of the question on the use of social networks and review websites as tools to support decision making during the travel was conducted, using a Principal Component Analysis (PCA). This analysis assists in associating and aggregating items whose variables explain a common concept. The aim is to determine whether the first two items (about social networks) will be grouped in a single group and the remaining ones (about review websites) in a different group, thus confirming that the sample behaves in a distinctively different way regarding each type of platform.

To this end, a factorial analysis was applied using the main components method and a Varimax rotation. Before factor retention, the suitability of the analysis was tested based on Kaiser-Meyer-Olkin (KMO) and Bartlett's Sphericity Test (see Table 5). In this regard, a KMO of 0.842 was obtained (a value which classifies the factor analysis as being of good quality), along with a significance of 0.000, thus rejecting the null hypothesis, which indicates that the correlation matrix is an identity matrix. In conclusion, enough conditions were met for the the PCA to be considered adequate.

Table 5. Bartlett's KMO and sphericity tests on applied PCA

Kaiser-Meyer-Olkin Measure		0.842
Bartlett's Sphericity Test	Approximate chi-square	1104.189
	Degree of freedom	15
	Significance	0.000

Source: Own Elaboration

As for the number of factors to be retained, all methods of selection suggest the retention of two factors. First of all, it is stated *a priori* that the six items of the question being studied can be segmented according to the type of platform: two of the items are about social networks and four are about review websites, forming two groups. Kaiser's criterion also suggests the retention of two factors, since these have an eigenvalue value which is greater than 1 (see Table 6); similarly, the Variance Percentage Criterion also points to the retention of two factors, since these together account for 85% of the variance - more than the 60% defined by the criterion.

Table 6. Total variance explained by the factors

Component	Initial eigenvalues	% of variance explained	% of cumulative explained variance
1	3,787	63 %	63%
2	1,284	21%	85%
3	0,413	7%	91%
4	0,254	4 %	96%
5	0,150	3%	98%
6	0,112	2%	100%

Source: Own Elaboration

The denomination of each of these factors is described in table 7, as well as the individual variables that compose them.

Table 7. Factors taken from the PCA and its component

Factor 1 - "Uses review websites as a tool to support decision making"
"I use review websites to read reviews from other users"
"I use review websites to see photos and descriptions of the space I'm looking for"
"I use review websites to look for additional information (timetables, location, prices, menus...)"
"I use reviews websites to find restaurants, lodgings, leisure spaces or others near me"
Factor 2 - "Uses social networks as a tool to support decision making"
"I use social networks to visit the spaces' institutional pages"
"I use social networks to ask for advice from my contacts"

Source: Own Elaboration

Table 8 shows the coefficients of reliability (Cronbach's Alpha) of the two factors, which vary between 0.95 and 0.73 for Factor 1 and Factor 2, respectively. These values, being higher than 0.7, suggest that there is internal consistency in each factor. In other words, one can say with relative certainty that each factor integrates questions that measure the same concept.

Still on the same table, Factor 1, about review websites, explains a greater part of the variance (63%), with only 21% of the variance explained by Factor 2. This circumstance may indicate that the respondents demonstrated a pattern of more easily identifiable behavior when responding to the items about review websites, which also justifies the higher degree of reliability in this factor, compared to that in Factor 2.

Table 8. Explanatory factorial analysis of the use of social media to search for information during the trip

| Factors	Loadings	Eigenvalue	Variance explained	Reliability
Uses review websites as a tool to support decision making		3.787	63%	0.950
"(...) read reviews of other users"	0.925			
"(...) see the photographs and descriptions of the space I'm looking for"	0.931			
"(...) seek additional information"	0.930			
"(...) discover restaurants, lodging, leisure spaces or others near me"	0.886			
Uses social networks as a tool to support decision making		1.284	21%	0.732
"(...) visit the spaces' institutional pages"	0.854			
"(...) ask for advice from my contacts"	0.896			

Source: Own Elaboration

4.3 Cluster Analysis

With the objectives of the present investigation in mind, it is pertinent to segment the study sample according to its social media behavior during and after the travel. For this purpose, a cluster analysis was applied, which facilitates the segmentation of respondents

into relatively homogeneous groups. The data extracted from the PCA was used to conduct this Cluster Analysis.

Firstly, a hierarchical cluster analysis was implemented, using the Ward aggregation criterion and the squared Euclidean distance as a measure of dissimilarity, on the standardized data. The results of this phase helped determine the number of clusters to be extracted in the analysis, through the subjective cut of the dendrogram, the elbow procedure (see Figure 1) and analysis of the fusion coefficients (see Table 9).

Figure 1. Cluster Analysis – Dendrogram and Line graph (elbow procedure)
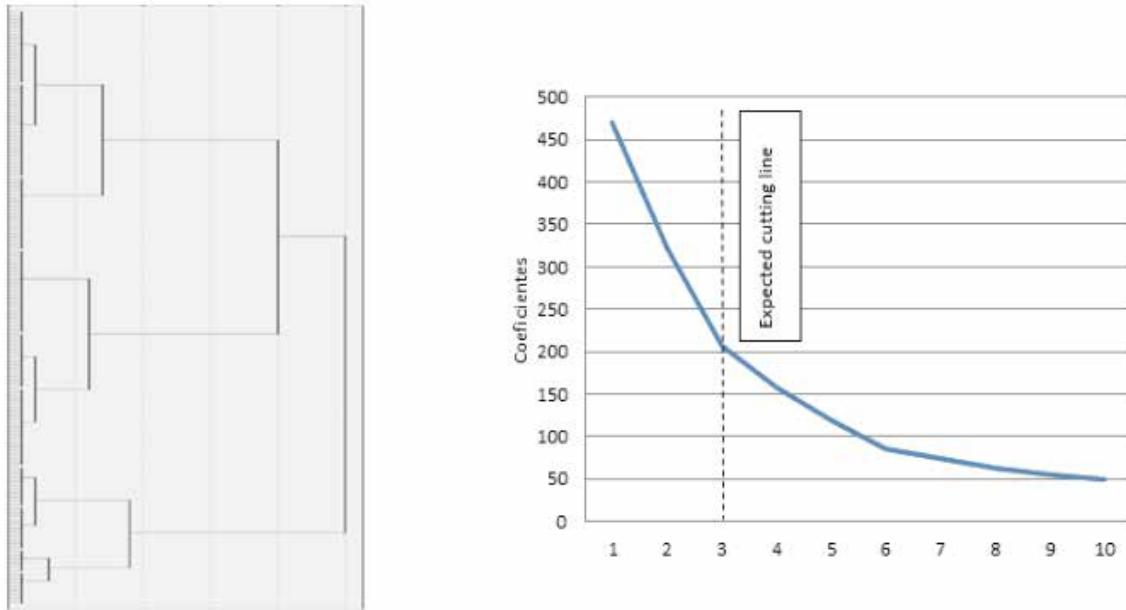
Source: Own Elaboration

Table 9. Cluster analysis – Table of fusion coefficients

Step	Cluster Grouped		Coefficients	Difference (coef. step N) – (coef. step N-1)	
	Cluster 1	Cluster 2			
226	5	7	49.649		
227	4	21	55.623	12.0%	
228	5	23	62.561	12.5%	
229	1	3	73.685	17.8%	
230	11	32	85.824	16.5%	
231	5	12	120.592	40.5%	
232	1	36	156.850	30.1%	
233	4	11	206.731	31.8%	3 clusters
234	1	5	322.871	56.2%	

Source: Own Elaboration

As all three methods suggested the same solution, three clusters were extracted: Cluster 1, with 95 individuals (40% of the 236 respondents who were considered for this analysis), Cluster 2, with 55 individuals (23%), and Cluster 3, with 86 individuals (37%).

It should be noted that all respondents who answered "No Answer" to any of the items in question 3.5 of the survey were not included in the cluster analysis, as they were initially excluded from the PCA. As such, the cluster analysis considered 236 respondents (94% of the total sample), while 14 respondents (6% of the sample) were not grouped into any of the clusters.

The most evident distinctions among the three clusters are based on their tendency to use certain types of social media as an alternative to others during the tourist trip, as well as on the propensity to use social media in general.

Essentially, Cluster 1 is composed of individuals who typically use review websites more frequently; on the other hand, Cluster 2 comprises individuals who preferentially use social networks; finally, Cluster 3 records the lowest use of any type of social media.

Additional analyzes allow for a characterization of each of the clusters. Table 10 shows the main sociodemographic distinctions between individuals in each cluster.

In terms of gender, all three clusters have a higher percentage of female respondents, but Cluster 2 stands out: while clusters 1 and 3 have nearly identical percentages on this regard, Cluster 2 shows an even greater predominance of female respondents (76%).

Table 10. Sociodemographic characterization by cluster

Variable	Cluster 1	Cluster 2	Cluster 3
Gender			
Male	35%	24%	34%
Female	65%	76%	66%
Professional situation			
Self-employed	4%	16%	13%
Employee on behalf of others	68%	66%	51%
Unemployed	3%	4%	4%
Student	23%	13%	26%
Retired	1%	2%	7%
Average monthly net income of the household			
Less than € 1500	32%	29%	29%
Between 1500 € and 3000 €	49%	49%	49%
Between 3001 € and 4500 €	16%	16%	16%
More than € 4500	2%	6%	6%
Average age	37 years	41 years	40 years

Source: Own Elaboration

As for the professional situation, Cluster 2 is the one with the lowest percentage of students (13%), most of whom are employed (whether on their own account or on behalf of others). Curiously, Cluster 3 is the one which simultaneously contains the greatest percentage of students (26%) and retired respondents (7%). Because this is also the cluster of respondents who less often uses social media during the trip as a means to support decision making, it is theorized that the students included in this cluster are not those who rarely use social media in general, but the who rely the least on these platforms to support decision making, since this was the specific criterion used in the segmentation of the sample. Finally, Cluster 1 is composed almost exclusively of employees (68%) and students (23%).

In terms of performance, the three clusters present fairly similar data, especially Cluster 2 and Cluster 3, whose values only differ by a decimal level. Cluster 1 has fewer individuals earning over € 4,500 per month than the other clusters and more individuals with incomes of less than € 1,500, although the differences are only around 3%.

Similarly, the average number of completed years of education and the average number of household members are approximately identical among the three clusters, thus, along with the average monthly net household income, these indicators are not ideal to differentiate the clusters from one another.

On the other hand, the variable of age presents some differences, albeit slight. Cluster 1 is the youngest, with an average age of 37 years old. Cluster 2, which contains the smallest percentage of students, is the oldest cluster, with an average age of 41 years old. The third cluster has an average age of 40 years old, very close to the average of the sample (39 years - see Table 2), which coincides with the simultaneously strong presence of students, employees and retirees in this group, as these suggest a greater age dispersion.

For a better understanding, however, it is particularly important to characterize each cluster according to their respective tourism profile and their respective perceptions and habits of social media consumption.

4.3.1 Social media - Perspectives regarding Internet usage while traveling

Firstly, the perspectives of each cluster on the access and use of the Internet while traveling were tested. The results (see Table 11) correspond to the average opinions of the individuals in each cluster, measured on a scale of 1 to 5, with 1 being "totally disagree" and 5 "totally agree" in relation to the statements presented in each paragraph.

Table 11. Perspectives of each cluster regarding Internet usage while traveling

Variable	Cluster 1	Cluster 2	Cluster 3
1. Not having access to the Internet while travelling is a good thing	2.2	1.9	2.2
2. It is convenient to carry at least one device with me that can connect to the Internet	4.3	4.7	4.2
3. I prioritize lodging options with free Internet access	3.9	3.9	3.7
4. When the weather is unfavorable, I spend more time on the Internet	2.4	2.9	2.7

Source: Own Elaboration

4.3.2 Social media - Frequency of use of social media

The frequency of use of social media represents a factor of high importance for the objectives of this study, since it is a way of measuring its prominence in the tourist experience of the individual. As such, to characterize the clusters in question, it is pertinent to analyze how often each group uses social media, both in their daily lives (see Table 12) and while traveling (see Table 13).

In a first analysis, it is clear that Cluster 1 is, out of the three clusters, the one which most often uses review websites, both in their daily lives and while traveling; in parallel, Cluster 2 makes the most use of social networks instead. Lastly, and in contrast with the other two clusters, Cluster 3 typically tends to use social media in general the least.

It is also interesting to note that Cluster 2 is the one which shows the smallest difference in behavior regarding the frequency of social media use while traveling, with only a slight decrease in the average value in social networks (from 4.0 to 3.7) and review websites (from 2.2 to 2.1). This behavior is very characteristic of this cluster, whose individuals are more adept at integrating social media in life in general and while traveling in particular.

Table 12. Frequency of social media use during the daily life by cluster

Social media	Cluster 1	Cluster 2	Cluster 3
Social Networks	**3.5**	**4.0**	**3.2**
Facebook	4.4	4.8	4.3
Instagram	2.6	3.2	2.0
Reviews Websites	**2.6**	**2.2**	**1.9**
TripAdvisor	3.0	2.4	2.0
Booking	3.2	2.8	2.4
Zomato	1.7	1.5	1.3

Source: Own Elaboration

Table 13. Frequency of social media use while traveling by cluster

Social media	Cluster 1	Cluster 2	Cluster 3
Social Networks	**2.9**	**3.7**	**2.4**
Facebook	3.5	4.5	3.2
Instagram	2.2	2.8	1.6
Reviews Websites	**2.3**	**2.1**	**1.5**
TripAdvisor	2.7	2.3	1.6
Booking	2.7	2.6	1.9
Zomato	1.5	1.5	1.1

Source: Own Elaboration

In contrast, the respondents in Cluster 3 are the ones that decrease their social media use the most, as this group can be characterized as individuals who naturally distance themselves from these platforms and make no use of them as tools to enhance the tourism experience.

Cluster 1 stays close to the total sample's average values in this regard, but its respondents stand out as being the most frequent users of review websites. This preference suggests that individuals in this group have a pragmatic view of social media: from their perspective, review websites in particular are a convenient way to obtain and share useful information with a global community of tourists, which can contribute towards bettering the travel experience.

4.3.3 Social media - Using social media to search for information while traveling

As this is the question used for dividing the sample into clusters (after the application of the PCA), it is of particular interest to ascertain the answers given by each cluster to each of its items. In the following table, the average answers are presented in a scale of 1 to 5, with 1 meaning "Never" and 5 "Always" (see Table 14).

Table 14. Use of social media as a tool to support decision making by cluster

Social media	Cluster 1	Cluster 2	Cluster 3
Social Networks	**2.2**	**4.2**	**2.0**
Consult official information	2.8	4.5	2.4
Ask contacts for advice	1.7	3.9	1.6
Reviews Websites	**4.2**	**3.5**	**2.2**
Read reviews from other travelers	4.2	3.5	2.2
View photos and descriptions of a space	4.3	3.5	2.3
See complementary information for a space	4.3	3.5	2.3
Discover new spaces in the surroundings	4.1	3.3	2.1

Source: Own Elaboration

Similar to the frequency of general use of social media, Cluster 2 continues to be the one that most uses these platforms, this time as tools to support decision making. However, confirming what was verified in the previous analysis, Cluster 1 shows a much more frequent use of review websites for information search than any other cluster (average of 4.2).

In social networks, Cluster 2 stands out as the one that most often uses these platforms to search for information during the trip, with an average of 4.2, compared to the average of 2.2 in Cluster 1 and 2.0 in Cluster 3.

4.3.4 Social media - Motivations that lead to self-expression on social networks

In order to understand what distinguishes the origin of the behaviors of each cluster on social media, the motivations that lead them to express themselves in social networks during the trip were analyzed (see Table 15).

As previously noted, Cluster 2 is the one which most favors the use of social network. Thus, it is unsurprising that this cluster shows the highest values regarding the various motivations presented in this topic. At the same time, Clusters 1 and 3, with lower social network usage rates, express very similar opinions to one another about their motivations for self-expression on social networks, suggesting that both clusters perceive these platforms in a similar way.

Table 15. Motivations that lead to self-expression in social networks per cluster

Variable	Cluster 1	Cluster 2	Cluster 3
Convenience	3.5	4.0	3.5
Show off	2.2	2.8	2.5
Share unique experiences	3.0	3.6	3.1
Invite others to live the experience	2.7	3.4	2.7
Share happy moments	3.4	3.8	3.4
Express sociability	2.3	3.1	2.3

Source: Own Elaboration

Nevertheless, it is important to note that the motivations "Convenience" and "Sharing happy moments" continue to have a very positive score throughout all clusters, confirming that these are universal motivations for all tourists, regardless of their position on the use of social media while traveling.

In addition, Cluster 2 is the only one which agrees, albeit with an almost neutral classification, that the desire to express sociability motivates its expression on social networks (3.1), distancing itself from the global sample's average for this item (2.5). This helps to describe Cluster 2 as individuals who consider the sharing of travel experiences on social networks to be a societal norm, as suggested by Fotis (2015). In other words, part of the reason why Cluster 2 shares their travel experience on social networks is because this is expected of any tourist and represents an opportunity to express sociability.

4.3.5 Social media - Nature of reviews published on review websites

The nature of the reviews that each cluster publishes on review websites was also studied (see Table 16).

This question also includes an item which clearly allows one to ascertain which individuals do not write any reviews on review websites.

Table 16. **Nature of reviews published on review websites by each cluster**

Variable	Cluster 1	Cluster 2	Cluster 3
Very positive and / or very negative ratings	17%	10%	5%
Positive ratings	8%	17%	17%
Negative ratings	0%	0%	1%
Ratings, regardless of classification	55%	42%	31%
Does not write reviews	21%	31%	46%

Source: Own Elaboration

Starting with the last paragraph, it is confirmed that Cluster 1 is the one that most often publishes reviews (79%), followed by Cluster 2 (69%) and, finally, Cluster 3 (54%).

Regarding the nature of these reviews, the results are in line with the general analysis of the sample, with the highest percentage of individuals from all clusters indicating that, when writing reviews, these can be either positive or negative.

Cluster 2 and cluster 3 are more likely to write only positive reviews (17%), compared to Cluster 1 (8%), the latter being more unbiased regarding the quality of the experience, also strengthening the notion that this cluster has a more utilitarian and pragmatic view of the Internet.

### 4.3.6 Tourist profile - Preferred holiday destinations

Based on the total number of tourist trips made by the sample in the years 2014 and/or 2015, this study analyzes the destinations visited by each cluster during their main vacations (see Figure 2, Figure 3 and Figure 4).

The results show that Cluster 2 is the one which most traveled in Portugal (55%) or in an EU country (43%), preferably Spain (15%). This may suggest that, preferring to travel short distances (and at a reduced cost), Cluster 2 is typically made up of domestic tourists or travelers who more often opt for lower investment trips. It can also be assumed that the respondents in Cluster 2 prefer to travel to a nearby destination so as to save money on transport and invest in superior accommodation.

Cluster 1 and Cluster 3 show very similar results, both with a slight majority of individuals preferring to travel outside Portugal on their main holidays and to very similar countries (mainly Spain, the United Kingdom and France).
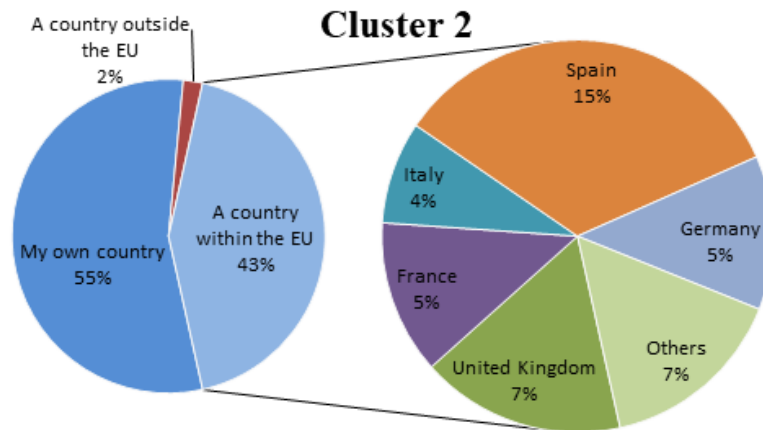
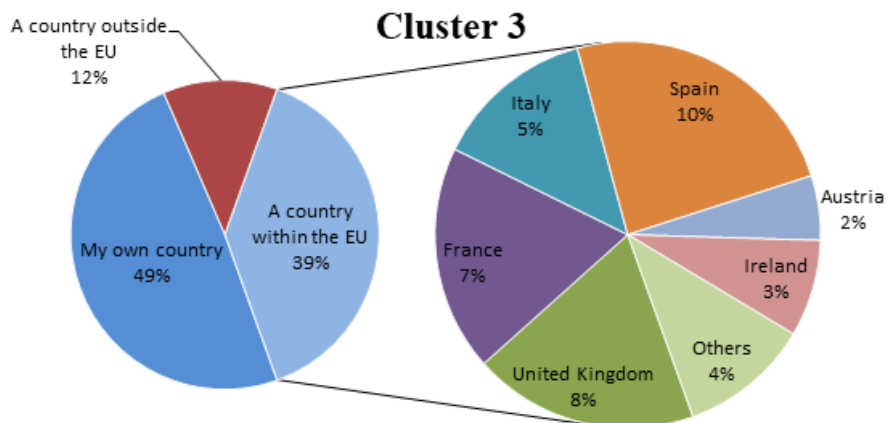**Figure 2. Destinations visited by each cluster on their main holidays (Cluster 1)**
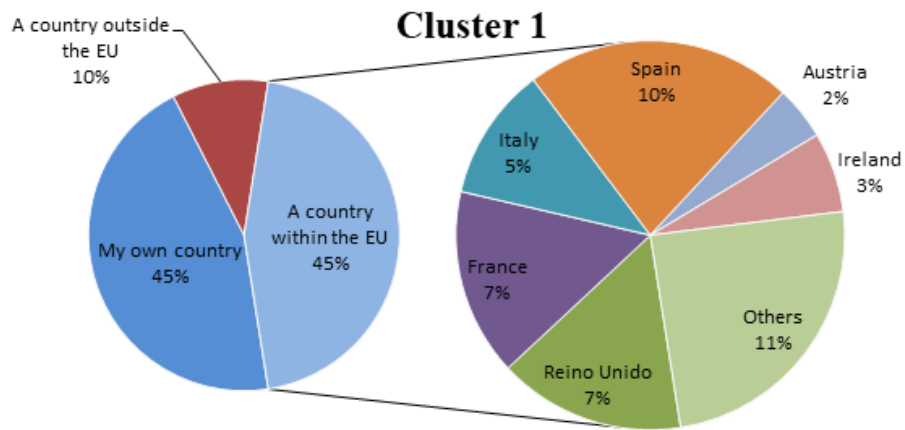
Source: Own Elaboration

Figure 3. Destinations visited by each cluster on their main holidays (Cluster 2)

Source: Own Elaboration

Figure 4. Destinations visited by each cluster on their main holidays (Cluster 3)

Source: Own Elaboration

5. CONCLUSION

In the case of social media, the presence of UGC is a basic element of these platforms. As such, given the steady growth in the number of social media users, the comments, opinions and experiences shared by tourists represent a substantial contribution to online tourist information. Additionally, the testimonies of other consumers are the source of information which tourists value the most when looking to make decisions during all stages of the trip, as confirmed by the literature review.

Through their study, Yoo, Lee and Gretzel (2007), Yoo and Gretzel (2008a) and Yoo and Gretzel (2008b) found that young males with high levels of income, education and internet ease-of-use and who travel more often are more likely to use social media to create content while traveling.

A questionnaire survey was implemented, allowing this study to work with primary data in two aspects: on the one hand, the validation of secondary data collected from other authors and, on the other, the identification of trends in the behavior of tourists on social media. This survey allowed for a better understanding of the behaviors and habits of social media use by Portuguese tourists who traveled in the years 2014 and/or 2015.

In the analysis of results, this study characterized the tourists that generate content on social media during the trip, noting that the results differ in part from those determined by Yoo, Lee and Gretzel (2007), Yoo and Gretzel (2008a), Yoo and Gretzel (2008b) and Yoo and Gretzel (2016). Gender and income level do not seem to have a significant influence on the propensity to generate content by Portuguese tourists, age seems to only partially influence this propensity, and the frequency of content generation only decreases after the age of 45. On the other hand, and validating the conclusions of these authors, age strongly influences the propensity to share photographs specifically, as young people are more predisposed to generate content in this format.

Through a Principal Components Analysis, the hypothesis was tested that, based only on the answers given by the respondents to the various items on the question about the use of social networks and review websites as tools to support decision-making during the trip, it would be possible to identify a distinct behavior in relation to the two items regarding social networks and the remaining four items regarding review websites. This hypothesis was confirmed, resulting in the extraction of two components: one which groups the variables on social networks and another which groups the variables on review websites. These components were used to segment the sample according to their response patterns, through a hierarchical cluster analysis.

The cluster analysis segmented the sample into three different clusters, which were thoroughly studied and characterized: Cluster 1, with about 40% of the 236 respondents admitted to this analysis, Cluster 2, with 23% of the respondents, and Cluster 3, with 37%. It was concluded that the main differences between the clusters are in the perception that they each have regarding social media and in the use that they make of them.

Cluster 1, which is also the most numerous, includes individuals who have a particularly utilitarian view of social media. In other words, for this group, social media (especially review websites) are a useful platform for obtaining and sharing information and keeping in touch with one's contacts.

For Cluster 2, social media (especially social networks) are also a source of entertainment and give the tourist the means to express sociability and communicate with their contacts through a more frequent sharing of travel contents. This preference for social networks is maintained even when searching for information, given that this segment favors social networks as a support for decision making while traveling.

Finally, Cluster 3, while not necessarily against the use of social media, is the least favorable towards these platforms, both in their daily lives and during the trip. In other words, this cluster is not particularly motivated to use either social networks or reviews websites, whether to generate content or to search for content generated by other users. However, results show extensively that this cluster does not necessarily have a negative opinion of social media.

At the sociodemographic level, although the three clusters do not differ substantially, Cluster 1 is the one that contains a greater percentage of male respondents, and is also the youngest group. Cluster 2, also the oldest group, is composed of the greatest percentage of female respondents and the smallest percentage of students. Cluster 3 contains both the highest percentage of students and retirees, reinforcing the notion that this is a very diverse group of individuals, not only demographically, but also in terms of tourism consumption habits. It is also added that the clusters present very similar results regarding "Average net monthly income of the household", "Number of household elements" and "Years of education completed successfully", so these indicators are not adequate to distinguish and characterize each segment.

The present study thus resulted in a comprehensive analysis of the behavior of Portuguese tourists on social media. This analysis was carried out in order to contribute to a better understanding of the motivations, process and implications regarding tourist generated content. In addition to the survey and analysis of empirical data, the literature review also represents a step towards the consolidation and aggregation of scientific literature in this area, which, given its relative youth, does not yet consist of an optimally robust body of publications.

In terms of future work the study will be continued to be developed, integrating Location Based Social Network Applications (LB-SNAs) that allows access to real-time tourist location information to investigate how this applications contribute and influence the touristic experiences and the usage behavior of tourists as also the problematics regarding privacy (Koohikamali, Gerhart & Mousavizadeh, 2015; Palos-Sanchez, Hernandez-Mogollon & Campon-Cerro, 2017).

ACKNOWLEDGEMENT

REFERENCES

Acar, A. S. & Polonsky, M. (2007). Online social networks and insights into marketing communications. *Journal of Internet Commerce,* 6(4), 55-72.

Adjei, M. T., Noble, S. M., & Noble, C. H. (2010). The influence of C2C communications in online brand communities on customer purchase behavior. *Journal of the Academy of Marketing Science*, 38(5), 634-653.

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.

Burgess, S., Sellitto, C., Cox, C., & Buultjens, J. (2011). Trust perceptions of online travel information by different content creators: Some social and legal implications. *Information Systems Frontiers*, 13(2), 221-235.

Carrera, P., Chiu, C-Y., Pratipwattanawong, P., Chienwattanasuk, S., Ahmad, S. F. S., & Murphy, J. (2008). My space, my friends, my customers. In P. O'Connor, W. Höpken & U. Gretzel (Eds.), *Information and Communication Technologies in Tourism 2008* (pp. 94-105). Vienna: Springer Verlag.

Charlesworth, A. (2009). The ascent of smartphone. *Engineering and Technology*, 4(3), 32-33.

Damian, R. (2014). *Understanding digital marketing: marketing strategies for engaging the digital generation*. Kogan Page Publishers.

Facebook (2016). *Company info*. Retrieved from http://newsroom.fb.com/company-info/

Fotis, J. N. (2015). *The use of social media and its impacts on consumer behaviour: the context of holiday travel.* (Dissertação de doutoramento não publicada). Bournemouth University, Reino Unido.

Gretzel, U., Lee, Y-J., Tussyadiah, I., & Fesenmaier, D. R. (2009). Recounting tourism experiences: the role of new media. *International Conference on Tourist Experiences: Meanings, Motivations, Behaviours*, 1-4 abril 2009, Preston, Reino Unido.

Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: Foundations and developments. *Electronic Markets*, 25(3), 179-188.

Instagram (2016). *Stats*. Retrieved from https://www.instagram.com/press/

Internet Live Stats (2016). *Internet Users in the world*. Retrieved from http://www.internetlivestats.com/

Jones, S. & Fox, S. (2009). *Generations online in 2009.* Pew Internet and American Life Project. Retrieved from http://www.pewinternet.org/Reports/2009/Generations-Online-in-2009.aspx

Kang, M. & Schuett, M. A. (2013). Determinants of sharing travel experiences in social media. *Journal of Travel and Tourism Marketing*, 30(1-2), 93-107.

Koohikamali, M., Gerhart, N., & Mousavizadeh, M. (2015). Location disclosure on LB-SNAs: The role of incentives on sharing behavior. *Decision Support Systems*, 71, 78-87.

Lenhart, A., Madden, M., Macgill, A.R., & Smith, A. (2007). *Teens and social media.* Pew Internet and American Life Project. Retrieved from http://www.pewinternet.org/Reports/2007/Teens-and-Social-Media.aspx

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), 458-468.

Munar, A. M. & Jacobsen, J. K. S. (2014). Motivations for sharing tourism experiences through social media. *Tourism management*, 43, 46-54.

Nielsenwire (2009). *Global advertising: Consumers trust real friends and virtual strangers the most*. Retrieved from http://blog.nielsen.com/nielsenwire/consumer/global-advertisingconsumers-trust-real-friends-and-virtualstrangers-the-most/

O'Connor, P. (2008). User-generated content and travel: A case study on Tripadvisor.com. *Information and communication technologies in tourism 2008*, 47-58.

Palos-Sanchez, P. R., Hernandez-Mogollon, J. M., & Campon-Cerro, A. M. (2017). The behavioral response to location based services: an examination of the influence of social and environmental benefits, and privacy. *Sustainability*, 9(11), 1988.

Pan, B., MacLaurin, T., & Crotts, J. C. (2007). Travel blogs and the implications for destination marketing. *Journal of Travel Research*, 46, 35-45.

Ryan, D. (2014). *Understanding digital marketing: Marketing strategies for engaging the digital generation*. Kogan Page Publishers.

Shao, G. (2009). Understanding the appeal of user-generated media: A uses and gratification perspective. *Internet Research*, 19(1), 7-25.

Tedjamulia, S. J., Dean, D. L., Olsen, D. R., & Albrecht, C. C. (2005). Motivating content contributions to online communities: Toward a more comprehensive theory. In *Proceedings of the 38th annual Hawaii international conference on system sciences* (pp. 193b-193b). IEEE.

Tnooz (2010). *How travellers use technology to search, book and play when away*. Retrieved from https://www.tnooz.com/article/survey-how-travellers-use-technology-to-search-book-and-play-when-away/

Van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content. *Media, culture, and society*, 31(1), 41.

Xiang, Z. & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism management*, 31(2), 179-188.

Yoo, K. H. & Gretzel, U. (2008a). Use and impact of online travel reviews. *Information and communication technologies in tourism 2008*, 35-46.

Yoo, K. H. & Gretzel, U. (2008b). Understanding differences between online travel review writers and non-writers. In *Proceedings of the 13th Annual Graduate Education and Student Research Conference in Hospitality and Tourism, Orlando, FL* (pp. 21-29).

Yoo, K. H. & Gretzel, U. (2008c). What motivates consumers to write online travel reviews? *Information Technology and Tourism*, 10(4), 283-295.

Yoo, K. H. & Gretzel, U. (2016). Use and creation of social media by travellers. *Social media in travel, tourism and hospitality: Theory, practice and cases*, 189.

Yoo, K. H., Lee, K. S., & Gretzel, U. (2007, Janeiro). The role of source characteristics in eWOM: What makes online travel reviewers credible and likeable. In *Proceedings of the 14th international ENTER conference in ljubljana, Slovenia* (pp. 23-34).

Yoo, K. H., Lee, Y., Gretzel, U., & Fesenmaier, D. R. (2009). Trust in travel-related consumer generated media. *Information and communication technologies in tourism 2009*, 49-59.