

Geolocalised social networks information: Can we trust it when designing cultural and tourist policies?

Víctor Temprano García (vtemprano@eco.uva.es), PhD Candidate Researcher of Marketing at University of Valladolid.

Rebeca San José Cabezudo (rebecasc@eco.uva.es), Associate Professor of Marketing at University of Valladolid.

Ana María Gutiérrez Arranz (anag@eco.uva.es), Associate Professor of Marketing at University of Valladolid.

María Redondo Carretero (mariareca@emp.uva.es), Adjunct Lecturer of Marketing at University of Valladolid.

Abstract

The aim of this research is to validate a new source of information based on geolocalised social networks, as a new approach to analyse the tourist's behaviour patterns on social networks.

How do we know tourists' preferences when visiting a city but they do not lodge in hotels, or they do not check in at any tourist office? The data gathered via geolocalised social networks, provides an innovative and valuable source of information to help policymakers to design cultural and tourist policies.

Using a software designed ad-hoc for these projects, we have been able to gather both quantitative and qualitative data from several geolocalized social networks such as Flickr, Picasa and Instagram, and we have analysed this data along with official data provided by the Valladolid City Council.

We have collected, catalogued and synthesised data, establishing a matrix to map the inputs to take into account, in order to design an efficient cultural and tourist policy in the city of Valladolid.

The preliminary findings pointed to a reliable data source, to be taken into account when designing and implementing cultural and tourist policies, along other validated information.

Key words: cultural and tourism policy making, geolocalized social networks effectiveness, social networks effectiveness, digital economics.

1. Introduction

More than 3,300 million people, some 46% of the world's population, are connected to the Internet, a figure which increases to around 88% in North America and some 73% in Europe (Internet World Stats, 2015). The number of mobile device users is also growing fast. Globally, there are 3,790 billion mobile phone users, which meant 51% world mobile device penetration by January 2016 (Kemp, 2016). In Spain, according to the latest data published by the National Institute of Statistics (INE) for 2015, 78.7% of households in Spain had access to internet, with the main type of Internet connection being through broadband established using a handheld device (smartphone or tablet). The mobile phone penetration rate in Spanish households was also reported to be 97%. According to the latest study published by the AIMC (AIMC, 2016), mobile phones are used by 93.9% of internet users to access the Internet, ahead of laptop computers (used

by 76%) and desktop computers (71.4%). Moreover, 77.2% of users access social networks every day.

Social networks are still primarily used for friendship purposes (the option chosen by 75.6% of users) and to keep up to date with current events (53.4%), family relations (49.2%), sharing hobbies (42.2%) and work relations (38.9%). Of the various social networks, Facebook, Youtube and Twitter are, in that order, the most widely used (IAB, 2015).

Both the high penetration rates as well as the use of Internet and the social networks, coupled with user willingness to create and share content, means that a huge amount of information is generated, in a wide variety of formats (written messages or texts, photographs, videos, etc.) and at breakneck speed. To take one example, according to an infographic published by the firm DOMO (James, 2012), by the time the Internet had reached the figure of 2,100 billion users, every minute over 680,000 items were shared, 100,000 tweets were sent, 48 hours of video were uploaded to YouTube and 3,125 new photographs were sent to Flickr, amongst the many other contents generated in other social networks and spaces on the Internet. The vast amount of information available coupled with the extraordinary speed at which it is generated in a variety of formats (texts, photographs, videos, etc.) provide the “seed” for the so-called Big Data, which has aroused so much interest today amongst scholars and practitioners. According to Gandomi and Haider (2015), *Volume*, *Variety*, and *Velocity* (or the *Three Vs*) has emerged as a common framework to describe big data (Chen et al., 2012; Kwon et al., 2014)

In this new landscape, organisations face the challenge of harnessing, storing and analysing all this information for decision-making purposes, and generating greater value for all stakeholders. Although the situation could be applied to any other organisation, our specific interest lies in ascertaining how local authorities can make the most of the potential offered by these new technologies and, particularly, social networks, in order to extract and examine the information generated by users with a view to devising more efficient strategies and policies. More specifically, our goal is to analyse the information of tourist interest generated both by residents and visitors to the city of Valladolid. Over the last few years, scholars and professionals alike have shown a great interest in the impact of social networks on the tourist industry, particularly vis-à-vis exploring its influence on tourist behaviour, either due to the importance of the communities of travellers to emerge on the Internet (e.g. TripAdvisor, Miguéns et al., 2008) or because of the influence which the information generated in social networks might have on the choice of tourist destination (e.g. Di Pietro et al., 2010). In the present research, we adopt a different approach and focus on examining the information from social networks which is used by public authorities to design more efficient tourist strategies and policies. Said interest stems from the project which a group of professors from the area of Marketing and Market Research at

the University of Valladolid¹ are currently working on in conjunction with Valladolid city council and of which the study now being presented forms part.

2. Experimental section

2.1. Social networks as a source of information

With developments in the Internet and web 2.0, information has ceased to be the preserve of a privileged few and is now distributed and managed by many thanks to the progress of information and communication technologies and, in particular, the Internet. Unlike conventional taxonomical or selective forms of distributing information, in the new 2.0 environment (reflected in phenomena such as blogs and, above all, social networks) said distribution is based on folksonomic or collaborative principles. In other words, all the members of the network participate and contribute towards distributing said information correctly, although not all do so equally, as there are veritable digital influencers (San José et al., 2012). Given such a context, beyond the actual information conveyed, it is individuals, acting as a node in these virtual networks, and their relations that are the real driving forces behind the scale and efficiency of the information's distribution.

Social networks are structures made up of people linked by one or more relations that may be friendship, work-related, hobbies or any other common interest they share. Social networks thus emerge as areas for social exchange, or conversational networks composed of individuals who play an active part by spontaneously sharing experiences and opinions on issues of common interest (Wiertz and de Ruyter, 2007). As a result of this exchange of experiences, members of the network are likely to influence one another in their decisions and behaviours (Kozinets, 2002). Social networks have expanded to embrace the phenomenon of user interaction, participation, and generated content. This engenders feelings of belonging and identification with the community, of sharing habits and traditions as well as a moral responsibility, thus leading to the creation of veritable communities (Muniz Jr and O'guinn, 2001; Wiertz and de Ruyter, 2007)

Social networks, or areas of online exchange (e.g., Facebook, Twitter, YouTube, etc.), have provided the seed for virtual communities, distinguishing them from physical communities (Jones, 1995), also known as online communities (Kozinets, 2002). Virtual communities constitute an excellent research context to explore individual behaviour and have attracted the interest of many marketing researchers (Dholakia et al., 2004; Porter, 2004). In these online spaces, individuals come together around some distinct interest to contact and interact with each other in order to exchange, share and pool resources, such as information, knowledge, experiences, entertainment, socio-emotional support and friendship, through diverse computer-mediated communication systems, and in a number of different formats (texts, photographs,

¹ Carmen Antón Martín, Carmen Camarero Izquierdo, María José Garrido Samaniego, Ana Gutiérrez Arranz (Study Director), Jesús Gutiérrez Cillán, María Redondo Carretero, Ana Isabel Rodríguez Escudero, Javier Rodríguez Pinto, Rebeca San José Cabezero, Chanthaly S.Phabmixay S.Phabmixay and Víctor Temprano García.

videos, etc.) (Jin et al., 2010). The communications which circulate around the communities have a great capacity to influence the attitudes and behaviour of the virtual community's members (Smith et al., 2005; Wangenheim and Bayón, 2004). This is particularly true of so-called influencers or individuals whose messages might reach far and gain substantial credibility. In sum, for De Valck et al. (2009) a virtual brand community is, first and foremost, an online community based on social communications and relationships. Internet applications may supply the relational online space, but member-generated content and interpersonal communication are the essence of a virtual community (Wu and Fang, 2010). Thus, the value of a virtual community comes from the contents produced and shared during member interaction and conversation (Jin, Lee y Cheung, 2010).

Ethnography is an anthropological method that has gained popularity in sociology, cultural studies, consumer research and a variety of other social science fields (Kozinets, 2002). Said research method involves observing the practices of groups of people and even being able to take part in them, so as compare what people say and what they do. When applied to consumer behaviour, "market-oriented ethnography" is an ethnographic focus on the behaviour of people constituting a market for a product or a service (Arnould and Wallendorf, 1994). In line with Chong (2010), adopting this market research technique is supported by the value it provides vis-à-vis gathering and recording data systematically on individuals' behaviour in a natural environment, thereby increasing the likelihood of spontaneously detecting moments or important answers in people's everyday lives. Moreover, being able to include multiple formats in the data observed (texts, photographs, etc.) can ensure the observer-researcher can enjoy greater reliability when interpreting behaviour.

"Netnography," or ethnography on the Internet (Kozinets, 1998), is a new qualitative research methodology that adapts ethnographic research techniques to the study of cultures and communities emerging through computer-mediated communications. As the advent of networked computing opens up new opportunities for market-oriented consumer interaction, so it also opens up opportunities for marketing researchers to study the tastes, desires and other needs of consumers interacting in online communities (Kozinets, 2002).

2.2. Example of a practical application: The case of Valladolid

The present study aims to gain an understanding of the behaviour of residents and tourists visiting the city of Valladolid based on the geolocalised photographs shared on social networks online as well as estimating the impact on tourism promotion of the photographs taken by individuals of Valladolid, information which the local authorities can employ to devise tourist policies that meet the needs of both groups of users. To achieve this, a process for observing and actively processing user generated information was set up based on the criteria pre-established by the research team.

Measures

In order to accomplish the above-stated objective, the research team listed a number of variables to be measured, related both to user profile –residents and tourists coming to the city of Valladolid- (for example their origin, number of photographs shared and the number of visits the photographs have received, i.e. number of views²) as well as to the actual photographs themselves (when and where they were taken) and shared (number of times the photographs were visited and the tags used to classify them). Said variables will provide information concerning which areas have sparked the greatest interest and when, both in situ and on the Net, in addition to being able to pinpoint the influencers who help to create the image that the city of Valladolid has on the Net, amongst other matters.

Data collection

Data were obtained on two photograph social networks: Flickr, a website for storing, organising and sharing photographs and videos and including geolocation tools for the places where the photographs were taken, indicating the exact location coordinates, and Picasaweb (Google), a digital platform for organising photographs that also includes text editing and photo editing tools. Table 1 provides a summary of the sources of information used and the key information taken from each for the purposes of the present study.

Table 1. Syntax of the database used and the information obtained.

Database	Fields	Information obtained
Flickr Picasa	Location	Visited place identification
	Position	User place of origin
	Distance between user and photograph	Type of user
	Views	Maximum, minimum and average views of a photograph taken in the city of Valladolid.
	Date a photograph was taken	Photographs filter based on the scope of time analysed.
	Category	Category of the most often photographed places
	Title and Tags (word associated to the photograph)	Social group generation by affinities
	Longitude A	Valladolid tourist areas mapping through geolocalised polygons. Own design. Visualization through ArcGis ³
	Latitude A	
	Longitude B	
Latitude B		

² See table 2 for a definition.

³ ArcGIS is an online cloud-based mapping platform available at: <https://www.arcgis.com/>

Results

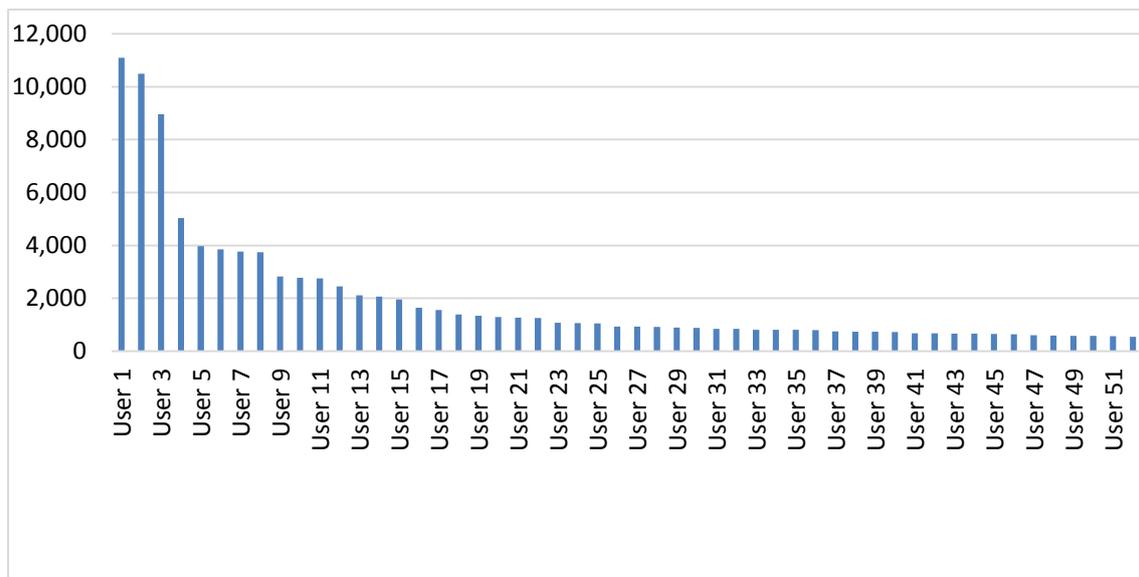
Having extracted and processed the data required for our study, we now present some of the results to emerge. Commencing with the information from users who shared images of the city of Valladolid in 2015 through the previously identified social networks (Flickr and Picasa), the final sample came to 226 individuals.

With regard to users' country of origin, 56.19% did not state their country of origin⁷. Of those who did, 40.27% are Spanish, while users from other countries (Italy, the UK, Australia, Colombia, Russia and Switzerland) accounted for a very small proportion, 3.54%.

As for the Spanish users, 99 users publicly declared their origin. In other words, we know where 43.80% of all users were from. Of these, 23.85% live in Valladolid, followed by users from Madrid, who account for 5.96%. The remainder come from the various other provinces in Spain (Barcelona, León, Santander, etc.).

As far as activity in the chosen social networks is concerned, users took and shared 8,426 photographs over the study period, with each photo being seen an average of 546 times, with a standard deviation of 1.352. Put differently, just a few users account for most visits to the photographs, as can be seen in Graph 1.

Graph 1. Distribution of users by the mean number of visits their photographs of the city of Valladolid received.



⁷ User origin is taken from the field "origen_user", which users fill in freely when registering on the Flickr and Picasa platforms. In line with current legislation concerning protection of personal details, our algorithm is designed to gather public information from users, edited by them.

Focusing more closely on the photographs, of the 8,426 taken, the 20 which aroused the greatest interest, judging by the number of times they were viewed, generated 152,720 visits. Most photographs were taken in March and September (respectively, representing when Holy Week and the local festivities in honour of the patron saint took place,) and accounted for 40% of all the photographs of the city of Valladolid uploaded to said social networks in 2015.

As regards the interest shown in the various areas of the city in 2015 (see Table 2), prominent are the Monuments, both in terms of the number of photographs taken of them as well as in the total number of visits to them. These are followed by the Squares and the Rivers of Light route. It should be pointed out that the Rivers of Light route feeds on and, in part, includes the other categories (Monuments, Squares, Church Monuments, Parks, Bridges, etc.). At the lower end of the Table are the markets, bridges and theatres. There are two possible explanations for this: on the one hand, these categories contain fewer items than the rest and, on the other, because in part they again take in other categories such as the previously mentioned Rivers of Light. As for the markets, it should be highlighted that Valladolid's most emblematic market, the Mercado del Vall, has been undergoing renovation work since October 2014.

Table 2. Distribution of the tourist areas in Valladolid with geolocalised images in 2015.

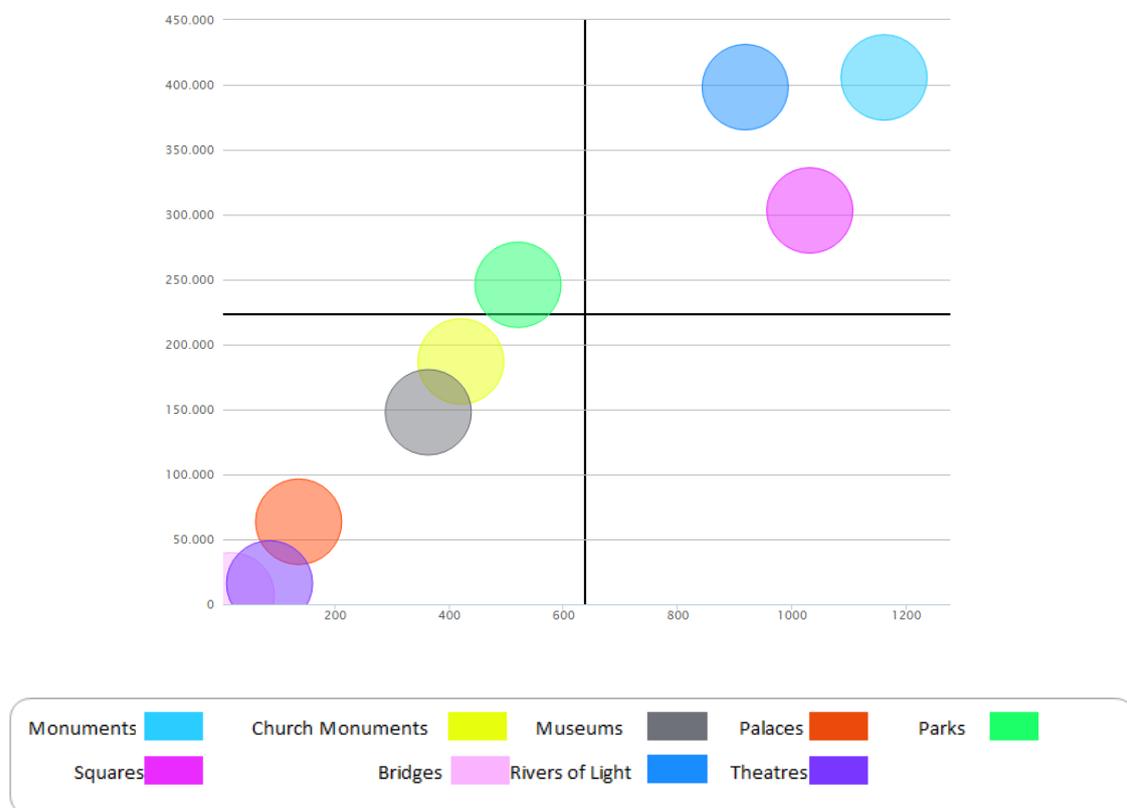
Category	Number of Photographs Frequency	Number of Photographs	Sum of views Frequency	Sum of views
Monuments	25,04 %	1.185	22,86 %	420.002
Squares	21,91 %	1.037	17,06 %	313.347
Rivers of Light	19,69 %	932	22,44 %	412.206
Parks	11,47 %	543	14,21 %	260.957
Ecclesiastic Monuments	8,92 %	422	10,22 %	187.728
Museums	7,71 %	365	8,08 %	148.417
Palaces	2,87 %	136	3,46 %	63.519
Theatres	1,80 %	85	0,87 %	16.037
Bridges	0,51 %	24	0,79 %	14.440
Markets	0,08 %	4	0,01 %	268
TOTAL		4.733		1.836.921

Image 3 shows the areas in the city that account for most of the geolocalised images. The tool used also allows us to analyse each area in greater detail. For instance, in one of the city's most emblematic spots, the Campo Grande park (Image 4), it can be seen how the images are concentrated at the main entrance, through Plaza Zorrilla, and, to a lesser extent, at the Plaza de

the average, (i.e. little interest on the part of the tourists) and the number of views of the pictures shared by users also below the average. The 2nd quadrant (bottom right) correspond to categories comprising items with the number of pictures taken above the average, (i.e. some interest on the part of the tourists) and the number of views of the pictures shared by users, below the average. The 3rd quadrant (upper left) correspond to categories comprising items with the number of pictures taken below the average, (i.e. little interest on the part of the tourists) but with the number of views of pictures shared by users, above the average. Finally, the 4th quadrant (upper right) correspond to categories that are comprised of items with the number of pictures taken above the average, (i.e. some interest on the part of the tourists) and the number of views of the pictures shared by users, also above the average.

As can be seen from Graph 2, the monuments hold the greatest interest for tourists, followed by the Rivers of Light route (which overlaps with several other categories in which it is included) and the Squares of Valladolid. Fewer photographs are taken of the parks during the year, although the photographs are seen quite a large number of times. The quadrant containing the fewest number of photographs seen and having the least impact includes Church Monuments, Museums, Palaces and, finally, Theatres and Bridges of Valladolid.

Graph 2. Matrix of tourist categories in the city of Valladolid in terms of popularity in 2015.

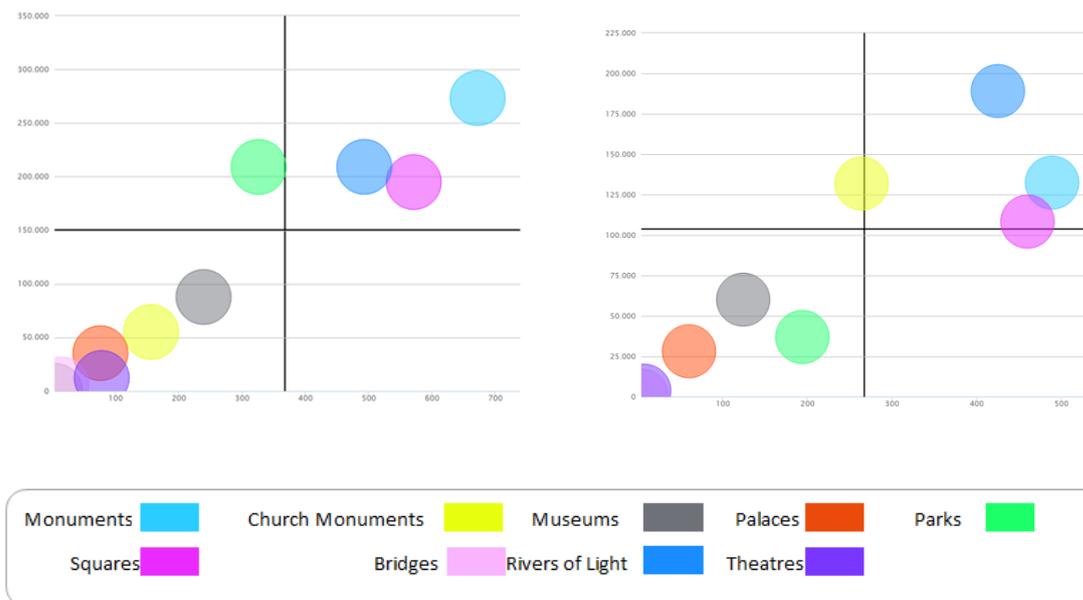


Although few data are available concerning user origin, we were keen to ascertain whether there were any differences apparent in the level of interest which the various categories aroused.

Residents showed a preference, in this order, for the Rivers of Light route, the Monuments, and the Squares. As far as the sample as a whole was concerned, greatest interest was shown for the Church Monuments. For their part, tourists evidenced a similar preference, albeit in a different order: Monuments, Squares and the Rivers of Light route, followed in this case by Parks (see Graphs 3 and 4).

Graph 3. Matrix of tourist categories of the city of Valladolid in terms of popularity in 2015. Residents.

Graph 4. Matrix of tourist categories of the city of Valladolid in terms of popularity in 2015. Tourists.



3. Discussion of results

It is an undeniable fact that Internet and social networks have become an integral part of people's lives. It is also no less true that organisations must take advantage of the opportunities afforded to them by the new technologies when it comes to decision making. People speak on social networks and organisations must listen if they are to devise valuable proposals that will meet the demands of today. Yet, this is by no means an easy task. The information generated in real time is so vast and comes from so many sources and in so many diverse formats (photographs, texts, etc.) that current data mining techniques prove insufficient. In this regard, the possibilities offered by ICT can help organisations to identify, capture, store and process said data, although this should be complemented by analysis and interpretation from those professionals most suited to the goals of the particular research in question.

In the present research, based on the research team's prior knowledge and following a previously established procedure, a tool has been designed to obtain data and process it. After analysis, results have emerged which we feel to be of considerable importance for a city's tourism management and image. Based on the spontaneous actions of private users (in which the observer has not intervened) in social networks, we have been able to identify which areas have aroused

the greatest interest amongst those visiting the city, on what dates and what influence the pictures of the city shared with other users might have had, amongst other aspects. We feel this information to be extremely valuable for local authority managers in that it helps them to pinpoint which areas need to be promoted further or exploited to a greater extent, or which need to be “improved”. The implications of this for management are clearly both wide-ranging and diverse: devising new routes or tourist package deals, using the areas which prove most appealing when conducting publicity campaigns for the city or how to carry out joint events between these areas and others, or geolocalisation based action (such as placing informative beacons that provide NFC signals) that would allow the visitor experience to be enriched, all of which would help to forge enhanced brand capital for the city. In this regard, identifying digital influencers and cooperating with them is of prime importance in that their actions may have a great impact on the scope and veracity of communications, helping to attract new tourists to the city. Local managers should implement policies that foster active citizen involvement.

Clearly, we must not overlook another of the main stakeholders involved in this study and who local managers equally seek to satisfy: local residents. They should also be listened to and actively engaged so as to make them ambassadors for the city. In this sense, encouraging a feeling of identity and belonging to the city is key.

On a different note, the information obtained might also help managers to evaluate the efficacy of the funding dedicated to promoting tourism or conservation, improving and maintaining the city’s heritage, in that these are measures which allow the efficacy of the action taken to be gauged in terms of user satisfaction.

Such is the scale of the phenomena being explored in our research that the present analysis may be viewed merely as the first part of the study we intend to carry out and, as such, it is by no means free from limitations. Firstly, with regard to how information has been gathered, the selection criteria followed to obtain information from one network or another is based on the possibility of extracting this information through the Application Program Interface (API) and the users that make use of this social network. An API is a set of routines, protocols, and tools for building software applications used to communicate with an operating system, databases or communications protocols.

Facebook has an API available but does not allow query parameters based information, events, terms, etc. Instagram has a very simple API, but the content is not based on tourism, therefore, is not acceptable for use in this study. Twitter allows queries and also has content, but the use of geolocation is decreasing. Flickr is a network specialized on photography with a large number of data geolocalised, and with an API for querying information based on content and geolocation, making it possible to analyse data used for the current research.

One further limitation is that the information accessed is public, in other words, it is the information which the user has authorised for open sharing. As a result, our tool fails to reflect data from users who have private profiles, that is to say, who share information with authorised users. Lastly, the final sample is made up of a large number of users who do not state where they are from. In future studies, we hope to be able to access larger and more representative samples that would enable us to conduct a more comprehensive and thorough statistical analysis.

With regard to how information has been gathered, the observer has neither participated nor interacted with the users observed. Although this might endow the results with greater objectivity and less bias, directly participating in the communities might help us to obtain another kind of information. In future stages of the research, we intend to delve more deeply into this issue. Furthermore, we feel that it would be timely to explore the possibility of cross-referencing the data obtained with that taken from other sources, whether primary or secondary (e.g. climate data) so as to gain a better understanding of the results to emerge and thus allow us to explore interesting new relations.

References

- AIMC (2016). Navegantes en la Red – 18ª Encuesta AIMC a usuarios de Internet. Retrieved from: <http://download.aimc.es/aimc/ROY76b/macro2015.pdf>. (Acceded, May 2016)
- Arnould, E. J., and Wallendorf, M. (1994). Market-oriented ethnography: interpretation building and marketing strategy formulation. *Journal of marketing research*, 484-504.
- Chen, H., Chiang, R. H., and Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165-1188.
- Chong, J. L. (2010). Evaluating the impact of Arnould and Wallendorf's (1994) market-oriented ethnography. *Journal of Business Research*, 63(12), 1295-1300.
- Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. (2004). A social influence model of consumer participation in network-and small-group-based virtual communities. *International Journal of Research in Marketing*, 21(3), 241-263.
- De Valck, K., Van Bruggen, G. H., and Wierenga, B. (2009). Virtual communities: A marketing perspective. *Decision Support Systems*, 47(3), 185-203.
- Di Pietro, L., Di Virgilio, F., and Pantano, E. (2012). Social network for the choice of tourist destination: attitude and behavioural intention. *Journal of Hospitality and Tourism Technology*, 3(1), 60-76.
- Gandomi, A. and Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35, 137-144.
- IAB (2015). VI Estudio Redes Sociales de IAB Spain. Retrieved from: <http://www.iabspain.net/wp-content/uploads/downloads/2015/01/Estudio-Anual-Redes-Sociales-2015.pdf>. (Acceded, May 2016)
- International World (2015). *Usage and population statistics*. Retrieved from: <http://www.internetworldstats.com/> (Acceded, May 2016).
- INE (2016). Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los hogares. Retrieved from:

http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=resultados&idp=1254735976608. (Acceded, May 2016)

- James, J. (2012). How Much Data is Created Every Minute? Retrieved from: <https://www.domo.com/blog/2012/06/how-much-data-is-created-every-minute/> (Acceded, May 2016)
- Jin, X.-L., Lee, M. K., and Cheung, C. M. (2010). Predicting continuance in online communities: model development and empirical test. *Behaviour & Information Technology*, 29(4), 383-394.
- Jones, S. (1995). Understanding community in the information age. *Sage, California*.
- Kemp, S. (2016). We Are Social [Press release]. Retrieved from: <http://wearesocial.com/specialreports/digital-in-2016>
- Kozinets, R. V. (1998). On netnography: Initial reflections on consumer research investigations of cyberculture. *Advances in consumer research*, 25(1), 366-371.
- Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of marketing research*, 39(1), 61-72.
- Kwon, O., Lee, N., and Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3), 387-394.
- Miguéns, J., Baggio, R., and Costa, C. (2008). Social media and tourism destinations: TripAdvisor case study. *Advances in Tourism Research*, 26(28), 1-6.
- Muniz Jr, A. M., and O'guinn, T. C. (2001). Brand community. *Journal of consumer research*, 27(4), 412-432.
- Porter, C. E. (2004). A typology of virtual communities: A multi-disciplinary foundation for future research. *Journal of Computer-Mediated Communication*, 10(1), Article 3.
- San José, R., Camarero, C., and Rodríguez, J. (2012). Looking for online opinion leaders: Why companies must identify and take care of more active users of online opinion spaces. *Universia Business Review*(35), 14-31.
- Smith, D., Menon, S., and Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of interactive marketing*, 19(3), 15-37.
- Wangenheim, F., and Bayón, T. (2004). The effect of word of mouth on services switching: Measurement and moderating variables. *European Journal of Marketing*, 38(9/10), 1173-1185.
- Wiertz, C., and de Ruyter, K. (2007). Beyond the call of duty: Why customers contribute to firm-hosted commercial online communities. *Organization studies*, 28(3), 347-376.
- Wu, S.-C., and Fang, W. (2010). The effect of consumer-to-consumer interactions on idea generation in virtual brand community relationships. *Technovation*, 30(11), 570-581.