



Cita bibliográfica: Yubero, C., Condeço-Melhorado, A. M., García-Hernández, M. y Catarina Fontes, A. (2021). Comparing spatial and content analysis of residents and tourists using Geotagged Social Media Data. The Historic Neighbourhood of Alfama (Lisbon), a case study. *Investigaciones Turísticas* (22), pp. 95-120. <https://doi.org/10.14198/INTURI2021.22.5>

Comparing spatial and content analysis of residents and tourists using Geotagged Social Media Data. The Historic Neighbourhood of Alfama (Lisbon), a case study

Análisis espacial y de contenido comparado entre residentes y turistas con el uso de datos geocalizados de redes sociales. El caso del barrio histórico de Alfama (Lisboa).

Claudia Yubero , Universidad Complutense de Madrid, España
claudiayubero@ucm.es

Ana Margarida Condeço-Melhorado , Universidad Complutense de Madrid, España
acondeco@ucm.es

María García-Hernández , Universidad Complutense de Madrid, España
mgarciah@ghis.ucm.es

Ana Catarina Fontes , Centre for Research and Studies in Sociology of Instituto
Universitário de Lisboa / Universidade NOVA de Lisboa, Portugal
Ana_Catarina_Fontes@iscte-iul.pt

ABSTRACT

Tourism flows to large cities have increased drastically in the past few years. The Alfama neighbourhood in Lisbon (Portugal) is facing major changes with respect to land uses, demographic features and social appropriation patterns in public spaces, caused by the intensification of tourism. The consequences of new emerging economic and symbolic values have rapidly given rise to a scenario of touristification and gentrification in the neighbourhood. In order to address such complexities, sustainable urban planning can benefit from real-time data sources that can represent the tourism flows in spatial and temporal perspectives. The research question allows Twitter to be used as an emerging data source and for analysing spatial patterns and content, based on two sample groups: residents and tourists, and their interpretations about the use of space for leisure activities. The research method is based on an analysis of two years of geotagged tweets in the city of Lisbon, differentiating between tourist and resident users, and, in a subsequent step, in the Alfama neighbourhood. The spatial distribution analysis and the content analysis have revealed not only spatio-temporal activity patterns but

also emotional responses to new trends in urban tourism uses, consumption and perception of an increasing tourism pressure in Alfama. The results are relevant in the field of tourism and sustainable urban planning.

Keywords: urban tourism; touristification; big data sources; Twitter; Alfama.

RESUMEN

Los flujos de turismo hacia las grandes ciudades han aumentado drásticamente en los últimos años. En Lisboa, el distrito de Alfama está experimentando un cambio funcional debido a los nuevos patrones turísticos y a la intensificación de su uso turístico. La creciente complejidad de la planificación urbana requiere fuentes de datos en tiempo real que puedan representar la temporalidad de los flujos turísticos. En este contexto, nuestra pregunta de investigación es ¿en qué medida el uso de nuevas fuentes de datos aporta nueva información sobre el uso del espacio para el ocio y el turismo, ya sea por residentes o turistas? El método de investigación se basa en un análisis de dos años de tweets georreferenciados, diferenciando entre usuarios turísticos y usuarios residentes en Lisboa. Un análisis de distribución espacial y un análisis de contenido han revelado no solo patrones de actividad espacio-temporales sino también respuestas emocionales a las nuevas tendencias en el uso del turismo urbano, el consumo y la percepción de la presión del turismo en Alfama. Los resultados proporcionan información importante para la planificación turística y urbana sostenible.

Palabras clave: turismo urbano; turistificación; fuentes de datos masivas; Twitter; Alfama.

I. INTRODUCTION

According to the World Travel Monitor report, international urban tourism increased by 16% in 2017 (Canalis, 2018) and “city breaks” reached 190 million trips (Messe Berlin, 2018). This growth has come with significant qualitative changes, including the sudden and disruptive irruption of new forms of urban tourism consumption and new business models in P2P platforms. Both aspects are closely related to an intensification in tourist use, leading to the over use of specific spaces within a city. In addition, the tourism footprint has extended well beyond the beaten track (Maitland & Newman, 2009). In the aftermath, urban sectors previously unaffected by the impacts of tourism have been transformed into tourist spaces. Indeed, tourism can be the trigger for accelerated processes of change, including profound transformations in the functional profile of urban spaces. Its impact can reach such a magnitude that it has also been considered a serious threat to the conservation of historic urban landscapes (HUL). The debate on overtourism, anti-tourism, touristification and tourism gentrification rages on the social media and on the academic community (Colomb & Novy, 2016; Wilson & Tallon, 2012).

These changing patterns of urban tourism entail urban planning challenges. The limited knowledge about contemporary tourism phenomena can be explained by the outdated traditional data sources, which are unable to provide sufficiently detailed information, on a time and space frame broad enough to cover the medium and/or long periods of time in which these processes operate. The inaccuracy of methods to study the patterns of tourist use at

the destination, operating nowadays urge for updated information that can support decision makers in implementing adequate sustainable policies integrating tourism and urban planning.

Lisbon is among the European capitals that have experienced the highest rates of growth in tourism and is considered one prominent case affected by the overtourism phenomenon (Milano, Novelli, & Cheer, 2019). The problem is exacerbated at the local level because tourism pressure is highest in specific places, including Alfama neighbourhood (Encalada, Boavida-Portugal, Ferreira, & Rocha, 2017). This neighbourhood is especially exposed to tourism gentrification due to its inherent features, combining the urban fabric and demographic aspects (Costa, 2008). In response to the pressure of change, its character is being diluted. The intensification of the tourist use is triggering profound transformations in the sociodemographic backbone of the community, that has been held accountable for a local and unique atmosphere imprinted in the narrow streets and alleys (Fontes, 2021). The case of Alfama is not exclusive in Lisbon, accelerated processes of touristification, are also affecting other historic neighbourhoods, such as Baixa and Bairro Alto (Malet Calvo, Gago, & Cocola-Gant, 2018; Gago, 2018, Mendes, 2017).

Such processes have been studied using research methods that address the phenomena from various disciplinary perspectives. However, regardless of the approach adopted, an analysis of tourist demand is crucial in order to understand and interpret the changes in urban spaces associated with tourism. Therefore, the integration of alternative tools and methods to reveal the ever more complex flows of visitors (volume, temporality and location) and patterns of tourist use at the destination (consumer trends) continues to be a priority in urban tourism research. The emergence of big data opened new opportunities and challenges for research (Shoval, 2018). They allow to push through traditional approaches based on “analogue” data collection techniques. Digital tracking technologies revolutionised the understanding and processing capacity of data, due to the possibility to access paramout data, harvested in a continuous time frame, resulting in huge volume and groundbreaking spatio-temporal accuracy for costumised geographic scales. Futhermore, big data enable information which can be useful in tracking and monitoring tourist behaviour. These interpretations derive not only from the geographical location but can also pursue emotional responses, based on a semantic analysis of messages on social media (Li, Xu, Tang, Wang, & Li, 2018; Marine-Roig, 2017; Shoval, 2018). Despite the already craved path to retrieve these sources and their inherent potential for the field of urban studies, particularly big data generated in social networks (Li et al., 2018, Salas-Olmedo, Moya-Gomez, Garcia-Palomares, & Gutierrez, 2018) there are still more applications to explore in tourism research.

In this context, our research question is to what extent the use of new data sources brings new information about the use of space for leisure either for residents or tourists? Taking an exploratory approach, the main goal is to test the viability of Twitter as a data source to illustrate the tourism behaviour of its users, whilst comparing residents and tourists in Alfama (Lisbon). The research method is based on an analysis of two years of geotagged tweets, differentiating between tourist and resident Twitter users in Lisbon. The data processing comprised a spatial distribution analysis and content analysis. The novelty of this methodology lies in

three aspects: 1) Combining a spatial and semantic analysis to reveal activity patterns and insights about Alfama's and Lisbon's tourism trends; 2) Twitter users were analysed separately while tracked as tourist or resident, in previous studies focusing urban tourism flows this topic has been considered a challenge subject to a high degree of uncertainty; 3) The use of content analysis has paved a way to interpret tourist use, retrieving patterns of tourist behaviour/consumption (e.g. resources, motivations, temporalities, opinions and attachment to places).

II. STATE OF THE ART. TOURISM PRESSURE AND BIG DATA SOURCES.

2.1. Tourism Pressure on Urban Spaces and Historic Centres

The growth of urban tourism in recent years has exacerbated the processes involved in the functional transformation of central urban spaces. These processes include population loss, increased tertiary uses, expansion of leisure activities and the disappearance of traditional commerce, in their aggravated stages they can be understood as gentrification and/or touristification processes/scenarios. The theoretical debate on the concept of gentrification has been around for a long time in academic studies (Clerval, Colomb, & Crieingen, 2011; Davidson & Lees, 2005; Glass, 1964; Hartman, Keating, & LeGates, 1982; Lees, 2008; Slater, 2009). However, only recently it has been contextualised for the field of tourism research (Wilson & Tallon, 2012), referring to a process of gentrification triggered by the growing tourism function in specific areas of post-industrial cities.

Tourism has been considered a vector of gentrification, regarding both production and consumption, creating staged and permanent new spaces prepared to answer the needs and expectations of affluent consumers, be they resident visitors or tourists (Cócola Gant, 2015; Gotham, 2005). Gentrified spaces, in which tourism and gentrification coexist, are produced for and consumed by a new cosmopolitan middle class that recreates similar urban environments wherever it goes (Judd, 2003). This definition implies one of the distinguishing features of the complexity of both processes: the practices of (new) residents and tourists are becoming increasingly similar. Indeed, some authors tend to agree that when we focus on behavioral aspects and corresponding practices of space, tourists and residents are becoming increasingly indistinguishable as groups of actors, whose specific intentions imprint, per se, contextual uses and practices in the city, hence working as triggers of change. On the one hand, the lifestyle of a resident tend to resemble that of a tourist, since a key element of urban development is the provision of cultural amenities and opportunities for consumption and leisure (Florida, 2002), understanding the city as an "entertainment machine" (Lloyd & Clark, 2001). On the other hand, according to post-tourism or third-generation tourism theories (Ashworth & Page, 2011; Hiernaux & González, 2014; Urry, 1990), the expectations and practices of a tourist tends to resemble that of a resident, since the tourist is a "resident on holiday". Hence, the practices of residents and visitors are alike and they are becoming increasingly interchangeable, to correspond to those of a new cosmopolitan, urban class with a medium-high economic level.

However, even if defining both groups can be considered a multi-layered task, other authors underpin the importance of understanding and highlighting the existing differences between both groups (Colomb & Novy, 2016). As for this study, we proposed a comparison of

spatial patterns and emotional response, by establishing a methodology to filter Twitter users in residents and tourists, aiming to contribute for this on-going discussion.

Most research about tourism gentrification has focused on the residential effects (Arias-Sans & Quaglieri-Domínguez, 2016; Colomb, 2012; Russo & Quaglieri-Domínguez, 2014) but groundbreaking work has also focused on its symbolic aspects (Sequera & Janoshka, 2013).

The term 'touristification' is another concept closely related to the growth and intensification of tourist use and refers to change in the functional profile of some urban sectors. At first, it alluded to the impact of the arrival of visitors and their role in transforming the economic and social fabrics of a territory. At any scale (from a neighbourhood to an island) they prompt a shift in the local economic-productive model, aiming for a focus on tourism activity by means of a highly specialised territorial model (Calle Vaquero, 2002). More recently, particularly due to the resounding message conveyed through the mass and social media, touristification has come to be used in reference to the negative impacts on residents, which comprise the loss or transformation of the facilities and services in their district to praise tourism. The concept of gentrification is at the root of touristification and it expresses to a certain extent the social rejection of tourism (anti-tourism). Thus, it is common to encounter headlines in the press mentioning "touristified spaces devoured by predatory tourism", often describing situations of high pressure caused by mass tourism (overtourism).

The intensification of tourist use in central urban areas is associated with new forms of urban tourism consumption that enlarge the tourism footprint, taking it beyond famous museums, landmarks and historic neighbourhoods. This progressive spread is encouraged by the availability of new facilities (associated for example with contemporary alternative cultural consumption and creative industries) in residential neighbourhoods, whose inhabitants are often specially vulnerable to gentrification processes. Consequently, the perception of change in territories previously out of the tourists's radar is becoming a concern. These historic and residential neighbourhoods are often vulnerable areas, which came to arise tourist's interest in their picturesque ways of life, as portraits of a more authentic experience and gateways of cultural immersion. The cases studied pile up, neighbourhoods in Berlin (Füller & Michel, 2014) and in Madrid (Lavapiés) (Sequera & Janoshka, 2013), or in coastal cities associated with sun, sea and sand tourism such as Palma de Mallorca (Morell, 2009), historic centres in emerging destinations such as Latin America (Hiernaux & González, 2014) and China (Liu, Zhu, Li, Sun, & Huang, 2019) or triggered by religious practices such as the Camino de Santiago (Pérez Guilarte & Lois González, 2018).

The growth of urban tourism flows, tourism gentrification, touristification, the overtourism and anti-tourism emerge as a series of interconnected processes. Other phenomena that would be worth mentioning as a catalyst and at the same time a consequence, in this context, is the outbreak of the P2P platforms for short rental accommodation such as Airbnb. They are held accountable for the rise of the land's price and rent, and therefore a cause for residential displacement (Cócola Gant, 2015) also in Lisbon (Mendes, 2020). In turn, this displacement prompts commercial and tourism gentrification and leads to the demise of urban infrastructures serving the local population (Kesar, Deželjin, & Bienenfeld, 2011), triggering

anti-tourism protest and resistance among residents, or what has come to be known as tourism phobia (Colomb & Novy, 2016).

2.2. Using Big Data to Analyse Tourism Pressure in Urban Spaces

Since 2010, traditional data sources (e.g. counts, surveys and interviews) have increasingly been complemented and/or replaced by big data (Batista & Silva et al., 2018). Tourism studies have been showing an increasing interest in the potentials of big data as sources of raw data and information (Baggio, 2016). The emotions of visitors or their spatial-temporal behaviour became questions whose answers can be disclosed through big data analysis, demonstrating a great potential for the discipline (Zhang, 2018).

This big data, valuable for tourism research, can be obtained from various sources, which can be divided into three categories: user-generated content (UGC), device-generated data and application-generated data (Li, Xu, Tang, Wang & Li, 2018). Many studies have employed UGC due to its availability, which means analysing data generated on social networks, such as Twitter (Bassolas, Lenormand, Tugores, Gonçalves & Ramasco, 2016). In fact, Twitter is one of the most frequently used sources (Murthy, 2013) because tweets provide free, real-time, geotagged information. Even if only 1% of tweets are accessed (Mahmud, Nichols & Drews, 2014), this actually provides a vast amount of data (around 500 million tweets are posted every day)¹. Valuable information on the tourism footprint can be extracted from tweets and sender profiles by examining movement patterns, the intensity of visits in specific places and content analysis.

There are also studies which favoured other UGC sources such as Panoramio (Bae & Yun, 2017; Encalada et al., 2017; García-Palomares, Gutiérrez, & Mínguez, 2015), Flickr (Crandall, Backstrom, Huttenlocher & Kleinberg, 2009; Donaire, Camprubí & Galí, 2014), TripAdvisor (Marine-Roig, 2017; Pantano, Priporas & Stylos, 2017), Facebook (Okazaki, Andreu, & Campo, 2017), Booking (Batista e Silva et al., 2018), Wikiloc (Lera I., Pérez T., Guerrero C., Eguíluz V., 2017) and travel blogs (Marine-Roig & Anton Clavé, 2015).

It has been demonstrated that big data can be used to shed light on the spatial tourism footprint thanks to the spatial-temporal accuracy of the generated data. The first step to enable geotagged data were personal GPS tracking devices. Processing the data generated can answer queries, while keeping the geographic location as the background to select samples of visitors (Birenboim, Reinau, Shoval, & Harder, 2015; Shoval, Schvimer, & Tamir, 2018; Zakrisson & Zillinger, 2012). This was subsequently replaced by the use of proxies such as data generated by mobile networks (Raun, Ahas, & Tiru, 2016) or user-generated geotagged photographs and messages posted on social networks such as Twitter (Hawelka et al., 2014), Weibo (Shao, Zhang, & Li, 2017; Shi, Zhao, & Chen, 2017), Flickr (Feick & Robertson, 2015; Girardin Fabien, Calabrese, Fiore, & Ratti, 2008), Panoramio (García-Palomares et al., 2015) and TripAdvisor (van der Zee, Bertocchi, & Vanneste, 2018) trading as Taylor & Francis Group The emergence of social media and Web 2.0 has a notable impact upon the tasks of destination managers

1. Source : <http://www.internetlivestats.com/twitter-statistics/>

as these platforms have developed into influential mechanisms affecting tourist behaviour. This paper shows how Destination Management Organizations (DMOs). These proxies have the advantage of maintaining a high spatio-temporal resolution while considerably increasing the data volume and currency and overcoming the possibility of bias (Shoval & Ahas, 2016).

Spatial distributions are considered a method to analyse tourists' spatio-temporal behaviour (Shoval & Ahas, 2016). The studies focusing on this kind of analysis identify hot-spots, where tourist interest and tourist use reaches a peak. The cities of Lisbon (Encalada et al., 2017) Vienna and Prague (Kádár, 2014) and Melbourne (Miah, Vu, Gammack, & McGrath, 2017) were showcased. Other study proposed a comparison between several European urban tourism destinations (García-Palomares et al., 2015). On a larger scale, more studies have been conducted, for example in Dresden (Hauthal & Burghardt, 2014), Hong Kong (Vu, Li, Law, & Ye, 2015) and worldwide, such as the study by Bassolas et al. (2016), who used Twitter to rank the world's most frequently visited sites. Another emerging line of research focuses on identifying the most visited tourist routes as a basis for recommendation and smart management at regional level (Chua, Servillo, Marcheggiani, & Moere, 2016; Vu et al., 2015) and urban scale (Comito, Falcone, & Talia, 2016; Manca, Boratto, Morell Roman, Martori i Gallissà, & Kaltenbrunner, 2017).

One of the methodological challenges that all these studies faced at some point, especially those using Twitter, is to identify the tourists whilst processing huge amounts of data. The given user's location, defined in his profile is insufficient because it is not always updated or not even specified. Studies that require higher precision have applied algorithms to identify the locations where the user is the most active, as it is recorded on social networks. Brogueira et al. (2016) used frequency to identify the regions in Portugal from which national tourists in the Algarve came, Salas Olmedo et al. (2018) defined tourists in Madrid as those who tweeted from there for less than one week within a year, while Manca et al. (2017) rised the time frame up to 20 days and García-Palomares et al. (2015) up to a month. These approaches have provided a more accurate portait of the tourism footprint, and have confirmed that visitors make a selective use of space. However, few studies to date have conducted a comparative analysis of the behaviour of residents and tourists (García-Palomares et al., 2015; Manca et al., 2017).

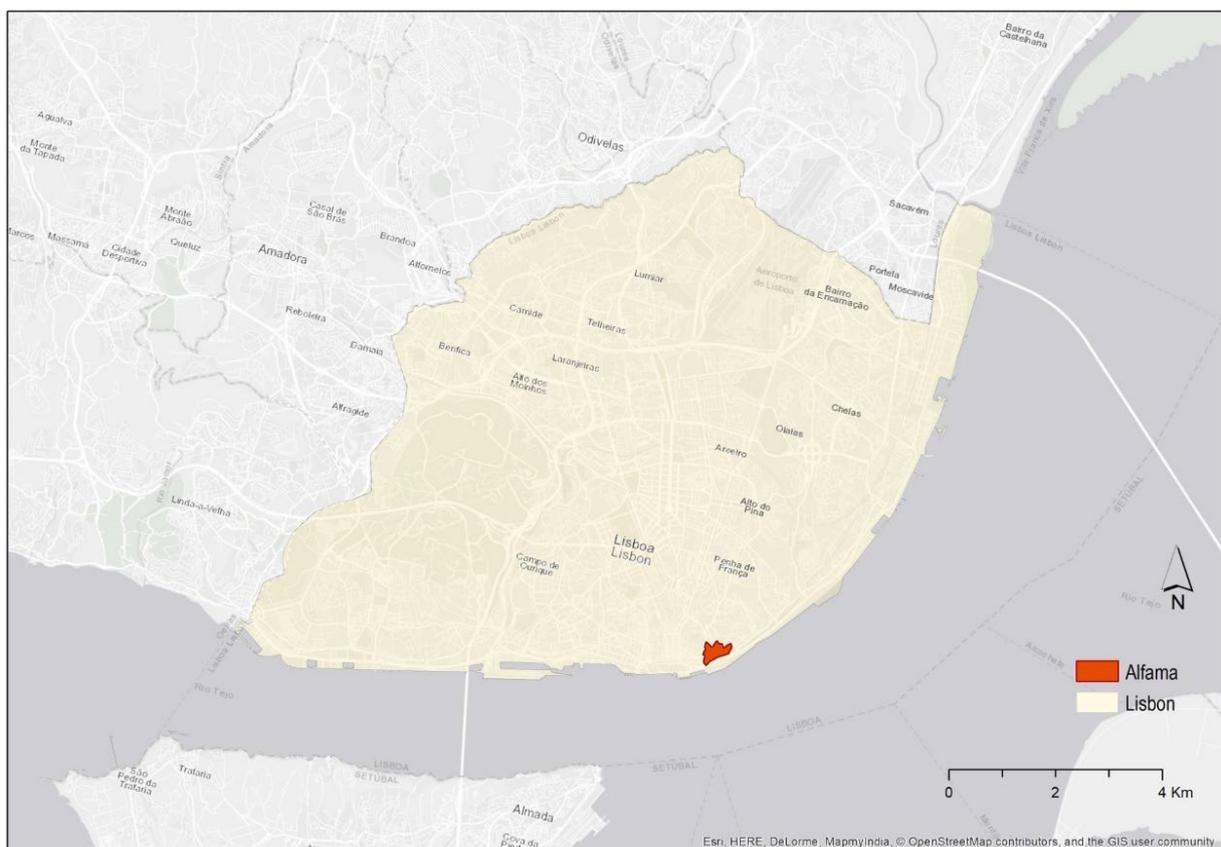
Alternative big data sources have most frequently been employed in tourism for content analysis (Li et al., 2018). The semantic and/or visual content associated with geotagged photographs has been used to determine perceptions of urban space (Girardin Fabien et al., 2008; Kim & Stepchenkova, 2015) or to identify the most frequently photographed element in a place (Hu et al., 2015; Miah et al., 2017). Twitter has also been used as a platform through which visitors and residents can report their satisfaction, opinions and experiences while travelling, analysed in terms of eWOM (Marine-Roig & Anton Clavé, 2015; Philander & Zhong, 2016; Sotiriadis & van Zyl, 2013; Williams, Inversini, Ferdinand, & Buhalis, 2017)000 relevant travel blogs and online travel reviews (OTRs. However, the connection between content and spatial analysis is less often found in the literature. The tweet content analysis has been used to determine a place of destination's image (Brogueira et al., 2016; Garay & Morales Pérez, 2017; Williams et al., 2017), which does not fully explore the potentials of spatial resolution.

In this theoretical and methodological context, designed to study touristification and the use of big data in tourism research, one of the challenges was tracking the spatio-temporal patterns of visitors (tourists and day-trippers), while combining spatio-temporal and content information. Additionally, a parallel study of tourist and residents' patterns of leisure activities provided useful information, when aiming for sustainable urban management in neighbourhoods already vulnerable and under pressure due to overexposure to tourism-recreational related flows and their impacts on site.

III. CASE STUDY. ALFAMA AND LISBON AS A TRENDING TOURISM DESTINATION.

The capital of Portugal, Lisbon, and the historic neighbourhood Alfama, are increasingly more popular tourism destinations. The two areas and also scales focused on the present study: the city and the neighbourhood are presented on the Figure 1.

Figure 1: Location of Lisbon and Alfama (Portugal)



Authors' own.

Lisbon is becoming ever more popular as a tourist destination, attracting a growing number of visitors. According to the *Global Destination Cities Index*, it ranked fifth among European cities with the highest rate of growth in tourism between 2009 and 2016, at 7.4% (Hedrick-Wong & Choong, 2016). This significant growth is confirmed by the statistics. The

number of travellers using Lisbon International Airport rose from 10.4 million in 2004 to 15.7 million in 2019. Meanwhile, according to the statistics, the port of Lisbon for cruise ships has also registered a rising affluence, from 305,185 passengers and 256 cruises in 2007 to 575,763 passengers 2019. Finally, the number of overnight stays in hotels in the municipality of Lisbon has more than doubled in the last years, rising from 6.5 million in 2010 to 18.6 million in 2019 (Statistics Portugal-National Institute of Statistics (INE)).

These figures result in the growth of tourist accommodation around the city. In 2009, there were 190 establishments in the municipality of Lisbon; by 2020, this number had risen to 237 (National Tourism Registry (RNT), a platform created by the Portuguese Tourism Board). Simultaneously, there has been a steep rise in the number of short-term rental (local accommodation (AL)), an option for tourism accommodation legally established in Portugal in 2008. The new legal framework permits the opening of hostels, which have multiplied since 2009 according to data from the National Registry of Local Accommodation (RNAL). In 2013, there were 54 registered establishments in Lisbon, but after the introduction of this new category (hostel), its popularity had made the term part of Lisbon's townscape. The number of apartments and rooms has also risen steeply, although doubts have been raised concerning the accuracy of figures obtained from the official register (RNAL) because many of them operate unregistered (Barata-Salgeiro, 2017). Nevertheless, the growth is indisputable: in December 2009 there were 42 registered local accommodation establishments (RNAL), a number which had soared to 12,315 by May 2018 and by October 2020 it had reached 19,404.

Tourist Lisbon experiences the emergence of tourist activities in the attempt to respond to this expanding market. Some are completely new to the city such as *Tuk Tuks*, others are based on a reinvention of local features and traditions, the Tram 28 (a journey by tram line 28 became one of the highlights of the "things to do" as a tourist in Lisbon, because of the classic tram's character going up and down the hills of the town and disclosing one historic neighbourhood after the other Graca, Alfama, Baixa, Chiado, Estrela) and the recreation of Portuguese cuisine and other local products (Henriques, 1996). All these activities have become part of everyday life and the urban landscape.

According to Benis (2011), in result of these processes the Alfama neighbourhood's character is fading away in shades that affect both the urban fabric and residents' way of life, in the trend capital and tourism are often prioritised over the preservation of local values (Gago, 2018; Fontes, 2019). Indeed, the increasing relevance of tourism has had major impacts on the real estate market, the situation has led to an imbalance of the access to long-term rental, while competing in a market that favours short-term rental. The regulation of such a balance has become one of the greatest challenges for urban management (Lestegás, Seixas, & Lois-González, 2019; Malet Calvo et al., 2018).

Given this panorama, criticisms of the growth of tourism in Lisbon are beginning to be voiced. According to a recent study published by the WTO (2018), a small percentage of inhabitants in eight European cities (including Lisbon) now consider that the development and marketing of tourism in their cities should be restrained. Malet, Gago and Cocola (2018) have reported a backlash against mass tourism in Lisbon led by artistic movements such as the Left Hand Rotation group, with their "Terremotourism" instructions, and political and social

associations (including local ones) and movements, demanding to take a step back in the tourism business whilst reclaiming residents' right to the city, especially regarding the right to housing (e.g. the Morar em Lisboa movement, the APPA in Alfama, the Rua dos Lagares

As a historic neighbourhood of Lisbon, Alfama is a magnet for urban tourism (Figure 2), and its symbolic and evocative power led it to represent the authentic picturesque side of the city (Cordeiro & Costa, 2003). Firmino da Costa (2008) has explored the question of the neighbourhood's cultural identity in terms of endogenous and exogenous factors, which can explain the relatively coherent narrative around its uniqueness and connection with resounding aspects of the "Portuguese culture" such as the Fado and summer street festivals (arraiais). Thanks to its historical background and cultural heritage, it is a paramount reference on tourism promotion and also as part of the collective imaginary of Lisbon. It recalls the restaurants offering traditional Portuguese cuisine, small bars for food drinks and Fado and Fado houses, and the Santos Populares festivities (Lisbon's celebration which spreads throughout the streets of the historic neighbourhoods of the city) (see Table 3).

Alfama still keeps its morphological features that reflect themselves the history of Lisbon since its foundation. The narrow winding streets, clinging to the hill located between the Tagus river and the castle, intersect and create ambiguous paths, difficult to apprehend in their logics of circulation. The studies of Benis (2011) and Gago (2018) describe the changes in tangible and intangible heritage as a consequence of a gentrification process, associated with the rehabilitation of buildings and public spaces, intensification in tourist uses and spread of tourist accommodation among the traditional residential area. This latter issue have escalated dramatically since the "local accommodation" (AL) was established and widely promoted on online platforms such as [booking.com](https://www.booking.com) and [airbnb.com](https://www.airbnb.com). In April 2018, there were 691 Airbnb advertisements for accommodation in Alfama.

Figure 2. Alfama neighbourhood as the conveyed image to promote tourism in Lisbon



Source: <https://www.mochileandoporelmundo.com/que-ver-y-que-hacer-en-lisboa-guia/>

IV. RESEARCH METHODS

4.1. Data Download and Processing

Twitter was selected because it provides free, real-time, geotagged information in large volumes. It also permits to carry out both a content and a spatial analysis. It is then seen as a valuable data source to analyse the impact of tourism and associated urban dynamics in Lisbon. From this source, the data was downloaded in real-time from the Twitter API over two years (from March 2016 to April 2018). The downloaded data generated a database with information stemmed from 339,511 Twitter messages (tweets) for Lisbon and 3,736 tweets for Alfama. This sample of users was expanded by downloading the last 3,200 tweets posted by each of them, for a total of 6,270,039 geotagged tweets. The area defined to perform the data download and the detailed neighbourhood study, takes into account the following criteria: in the case of Lisbon, a rectangle that approximately covered the entire municipality was used, as for Alfama, the area corresponds to the limits used by [booking.com](https://www.booking.com) platform to define the accommodation offered in this area².

Each of the tweets contained data on the X, Y coordinates, user ID, the date and time the tweet was posted, the device language setting and the text of the tweet, as well as other data of less interest for the present study. The data were entered into a geographic information system (ArcGIS 10.4.1) to create a layer of points based on the X and Y coordinates of the position from which each tweet was posted.

Subsequently, the data were filtered to eliminate users that might cause noise in the analysis, such as businesses and bots. Thus, we eliminated users posting a mean of more than 10 tweets a day who did not move (i.e. who tweeted from a single location). In addition, to determine the activity patterns of residents and tourists, it was necessary to analyse users rather than tweets. To this end, tweets were aggregated spatially and temporally (every 15 min) according to user ID, to obtain unique users in each spatial unit. Spatial aggregation was performed using a mesh of 400 m diameter hexagons, which had the advantage of mitigating the modifiable spatial unit problem (Openshaw, 1984), rather than using administrative divisions of different sizes, such as parishes.

Lastly, each user's place of residence was inferred by contrasting their tweet history with a world map layer and determining the country with the highest frequency of tweets for each user. This enabled us to distinguish between users residing in Portugal and those residing abroad. For Portuguese users, we also compared their tweet history with a NUTS-3 layer to determine their possible region of residence. This enabled us to divide the study sample of tweets into two broad categories: tweets posted by residents of Lisbon and its metropolitan area and tweets posted by tourists (foreign and national).

2. We used the definition applied on this platform because the Alfama district does not comprise an administrative division. Nevertheless, the heart of Alfama has been consistently considered the Santo Estevão and San Miguel parishes. This comprises an area where the urban fabric still reflects the medieval legacy (Firmino da Costa, 2008; Vieira da Silva, 1930,1943).

These two categories were contrasted with data on the content of texts associated with geotagged tweets posted in Alfama (3,736). We conducted a first exploratory analysis using keyword filters and then performed a detailed reading of the tweets to define categories and systematise the information. Text messages on Twitter are very short, comprising a sentence or several keywords describing a photograph. They also contain an abundance of personal comments (moods, greetings, etc.) and references to geographical location. Consequently, we created two categories of tags: location and content. The location category included 7 variables for analysis: 1. Lisbon, 2. Alfama, 3. Street/square/alley, 4. Restaurants, 5. Viewpoints, 6. Portugal and 7. Undefined. These tags were assigned based on the most specific location. For example, given a reference such as “Miradouro de Santa Luzia - Lisbon - Portugal”, the spatial reference selected was the St. Luzia viewpoint. As for the content category, we used tags that would enable us to generalise trending topics in the comments, including 1. History, 2. Architecture/landscape, 3. Atmosphere, 4. Fado, 5. Urban art/street art, 6. Restaurants/bars/meals, 7. Santos Populares (Lisbon festivities), 8. Accommodation, 9. Personal comments, 10. ISPA (University), 11. WebSummit (event), 12. Tram 28, 13. Shopping, 14. Publicity, 15. Trips/festivities/visits, 16. Art/galleries/exhibitions/museums, and 17. Complaints. These content tags were based on an assessment of the data, fieldwork in Alfama and also took into account literature review (Encalada et al., 2017; Henriques, 1996).

4.2. Analysis of Users’ Spatio-Temporal Distribution and Semantic Content of Tweets

Spatio-temporal disaggregation was performed by segmenting users, classified as residents or tourists. For each of these categories, we conducted a graphical analysis of their daily activity, comparing a typical working day (Tuesday) with a weekend day (Saturday). We also compared mean activity in Alfama and Lisbon. A spatial analysis was performed using density maps and descriptive statistics for the total number of users in each hexagon. To determine places with a statistically significant concentration of users, we used the local indicator of spatial autocorrelation (LISA), which analyses the distribution of a variable’s values in one location in relation to the values recorded in neighbouring locations (Anselin, 1995). This analysis identifies High-High clusters when a location presents a high concentration of users and is at the same time surrounded by other locations with similar characteristics. Low-Low clusters comprise the reverse situation. It is also possible to identify outliers such as High-Low when a location presenting a high concentration of users is surrounded by less frequented locations and Low-High locations representing the opposite case. For this statistical analysis, we used the inverse distance weighting (IDW) method and a previously calibrated neighbour distance of 1440 m. The analysis was performed comparing the distribution of residents and tourists, and the results were subsequently combined to classify areas in Lisbon according to the presence of each type of user. This enabled us to distinguish between markedly residential areas, tourist areas or mixed-use areas.

The spatial analysis proved very satisfactory for Lisbon as a whole, providing a very accurate location of tweets (with an estimated error of approx. 10 metres) and enabled us to situate Alfama in the context of the city (in comparison with other districts in relation to the greater or lesser presence of tourist Twitter users). Furthermore, for the more detailed scale

(Alfama), we conducted a content analysis, aiming to interpret processes of intensification of specific uses associated with the activation of tourist resources. Based on the two categories of users (residents of Lisbon and tourists in Lisbon), we also conducted a descriptive analysis of simple frequencies of content tags. The linear discriminant analysis (LDA) method was applied to the data sample.

V. RESULTS

5.1. Provenance of National and International Tourists in the Study Sample

As mentioned above, the data sample for the entire study area was used to segment Twitter users (48.046 in total) who had posted geotagged tweets, according to their country of residence. Portuguese residents comprised the largest group, followed by residents of Spain, the United States, the United Kingdom, Brazil and France (Table 1). Other countries among the top 10 were Italy, Germany, the Netherlands and Canada. Residents of the metropolitan area of Lisbon accounted for 91% of Portuguese residents. The remaining 9%, categorised as national tourists, came from other urban areas in Portugal such as the metropolitan areas of Porto, Coimbra and the Algarve. Since these findings point to an insufficient representation of national tourism (residents in other regions of the country), our results will distinguish solely between residents and tourists, without pursuing further data about international and national tourists.

Table 1. Country / region of residence of Twitter users in the data sample

Countries	Number of users	Number of users (%)	Portuguese NUTS 2 regions*	Number of users	Number of users (% Portugal)
Portugal	22965	47,8	Lisbon metropolitan area	21011	91,5
Spain	4679	9,7	Porto metropolitan area	587	2,6
USA	4327	9,0	Coimbra regions	204	0,9
United Kingdom	3310	6,9	Algarve	192	0,8
Brasil	2777	5,8	West region	167	0,7
France	1504	3,1	Aveiro regions	106	0,5
Italy	1165	2,4	Madeira autonomous region	91	0,4
Germany	808	1,7	Leiria region	75	0,3
Netherlands	504	1,0	Cávado	74	0,3
Canada	492	1,0	Azores autonomous region	62	0,3
Others	5515	11,5	Others	396	1,7
TOTAL	48046	100	TOTAL	22965	100

Authors' own.

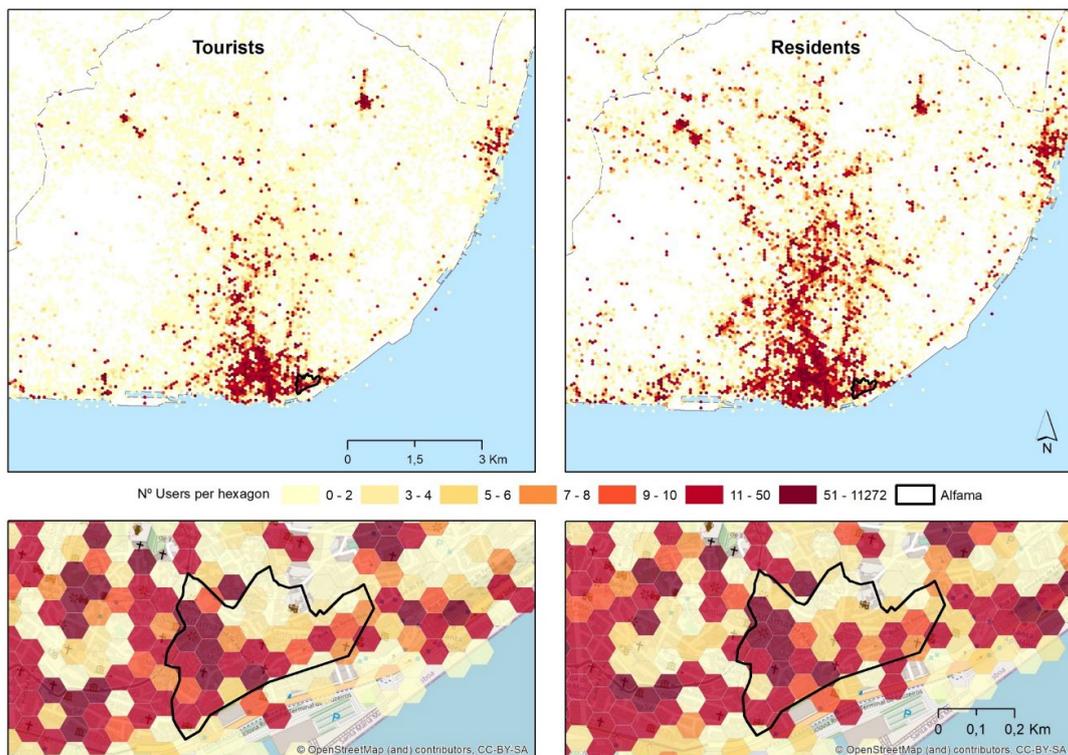
5.2 Spatial Density of Tourists and Residents in Lisbon and Alfama

The map in Figure 3 shows the spatial density of Twitter users in Lisbon, and in a further step zooming in on Alfama. At a city scale, the spatial distribution of tourists was less scattered than that of residents. Tourists were mainly concentrated in the central spaces of Lisbon, such as the Praça do Comércio, Baixa, Chiado, Bairro Alto, Cais do Sodré, the Avenida da Liberdade and the Marquês do Pombal area, also central for the residents' daily life in the city. Tourists also congregated in Alfama, in some parts outnumbering the residents.

Even if the results are less accurate when we zoom in on the neighbourhood, we can conclude that the places where tourists and residents gather the most correspond to the two most important squares within the neighbourhood, the Largo de São Miguel and Largo do Chafariz de Dentro. These results are coherent with the use of space observed on site, as the two squares are paramount for social life and the core from where or to where the paths through and in Alfama lead or initiate. Furthermore, when we consider the street network and implicit hierarchies, squares are essentially places of congregation, a potential endpoint for streets and a breakingpoint for strictly functional pedestrian circulation (Fontes, 2021).

We can also consider an highlighted axis in the maps, outlining the connection to the viewpoints on the upper part of the neighbourhood, and actually at its border. The Miradouro das Portas do Sol which somehow merges with the Miradouro de Santa Luzia in a series of platforms, are a favourite point to overlook the river and the neighbourhood.

Figure 3. Spatial distribution of Twitter users in Lisbon and Alfama. Tourists (maps on the left), residents (maps on the right)

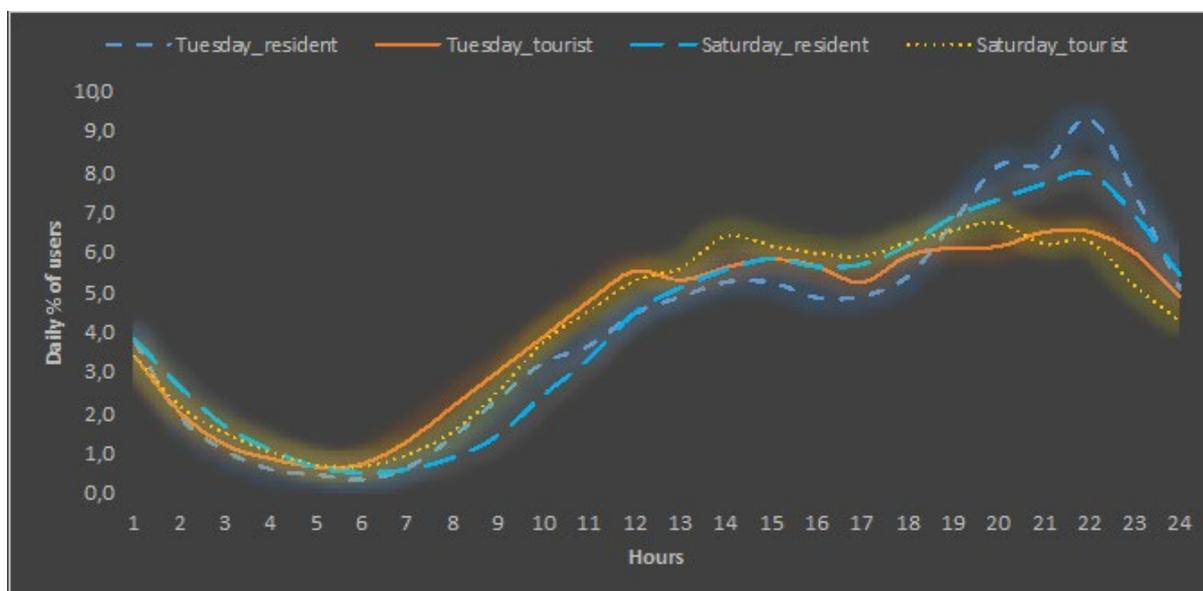


Authors' own.

5.3. Temporal Distribution of Tourists and Residents in Lisbon and Alfama

An analysis of the daily activity of tourists (orange) and residents (blue) on a typical working day (Tuesday) and a weekend day (Saturday) is generally similar (Figure 4). The main differences can be found on working days, where residents presented an increase in activity at 7-8 am, and two clear peaks in activity at around 2 pm and between 8 pm and 10 pm, the latter being more pronounced, with 9% of active users. These peaks coincided with out-of-work hours and therefore with the existence of more free time to access social media. At the weekend, residents showed a similar pattern, except for greater activity in the evening and less activity at night. Generally, tourists presented more stable Twitter activity throughout the day. The cause would be tourists have more free time to access social media and therefore their activity is more evenly distributed. Another remarkable difference was the decline in tourists' activity at 7 pm, which coincided with the residents' peak hours of activity.

Figure 4. Daily temporal distribution of geotagged tweets in Lisbon



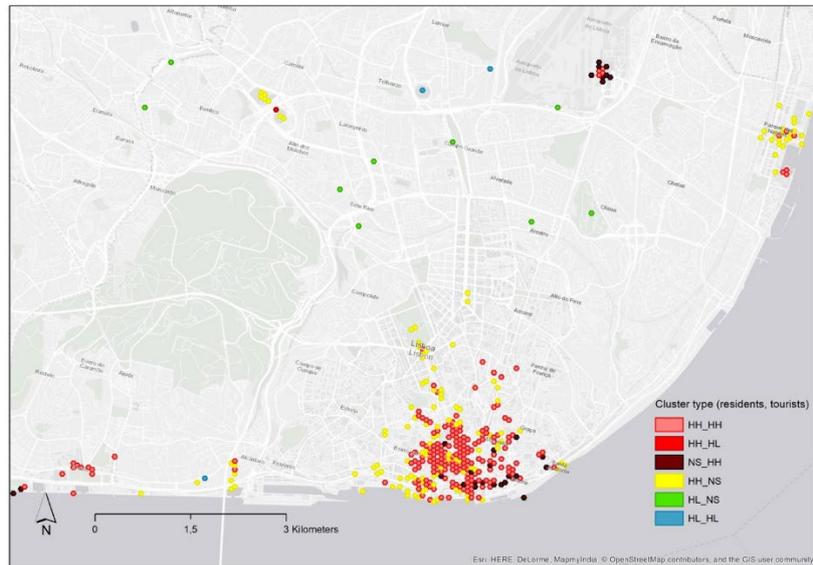
Authors' own.

5.4. Tourists' and Residents' Spatio-Temporal Patterns. A Comparison Between Lisbon and Alfama

The local spatial autocorrelation analysis was aiming to determine those areas of the city presenting a statistically significant concentration of Twitter users. Combining the results of this analysis, conducted for residents and tourists, revealed the existence of mixed-use areas characterised by a high concentration for both resident and tourist Twitter users (HH-HH), especially in the city centre (Figure 5). Some areas showed a very clear tourist specialisation (NS-HH), mainly in the lower part of the city, Alfama and Baixa. Other areas presented a dominant residential use, such as the Parque das Nações area. Unique elements in the city such as the cable car, the zoo and the universities presented highly concentrated clusters of Twitter users surrounded by areas of low concentration in the case of residents (HL-NS), but this was

not significant for tourists. Other elements, such as football stadiums, appeared as isolated centres of activity for residents and tourists alike (HL-HL).

Figure 5. Classification of areas in Lisbon according to the presence of residents and tourists



Authors' own.

Alfama's tourism-recreation specialisation compared with the rest of the city was also reflected in the percentages of tourists (Table 2). In Alfama, these accounted for almost 66% of Twitter users, 10% more than for the city as a whole. Geotagged tweets in Alfama posted by residents indicated that Lisbon's population also used this area regularly for leisure and recreation.

Table 2. Users (residents and tourists) in Lisbon and Alfama

Type of user	Lisbon	Alfama
Residents	20 652	779
Tourists	26 859	1 506
Total users	47 511	2 285
% Tourists	56.5	65.9

Authors' own.

5.5. Tourists and Residents in Alfama: Tweet Content Analysis

A detailed analysis of the tags assigned to the content of the 3,736 tweets in the Alfama sample is coherent with the previous conclusions in which the neighbourhood was facing a process of specialisation triggered by tourism. Although, many tweets simply referred to the location (18.6% of tagged comments, containing messages such as "I'm here, in Lisbon, in Alfama, at such and such a viewpoint") or contained personal comments (23%), others

contained more specific thematic content (Table 3). The most frequent of these latter referred to aspects related to the landscape and architecture (20% of the thematic comments), the context of the trip or visit (16%), bars and restaurants (14%), the atmosphere (9%) and Fado (7%), and these percentages were higher for tourists than for residents.

Table 3. Alfama. Thematic content of tweets (multiple response) in the data sample

	Resident		Tourist		Total		LDA
	Total	%	Total	%	Total	%	
Architecture/landscape	57	11.47	460	21.89	517	19.90	0.95
Trips/visits	22	4.43	392	18.66	414	15.94	1.85
Restaurants/bars/food/drink	50	10.06	321	15.28	371	14.28	1.04
Atmosphere	33	6.64	208	9.90	241	9.28	0.86
Fado	53	10.66	126	6.00	179	6.89	0.17
Publicity	155	31.19	16	0.76	171	6.58	-4.73
Santos Populares (Lisbon festivities)	25	5.03	113	5.38	138	5.31	-0.41
Urban art/street art	9	1.81	106	5.05	115	4.43	1.20
WebSummit (event)	3	0.60	86	4.09	89	3.43	1.83
Art/galleries/exhibitions /museums	10	2.01	62	2.95	72	2.77	0.65
History	4	0.80	63	3.00	67	2.58	1.17
Accommodation	13	2.62	54	2.57	67	2.58	1.04
Tram 28	4	0.80	58	2.76	62	2.39	1.37
ISPA (University)	55	11.07	0	0.00	55	2.12	-4.58
Shopping	2	0.40	28	1.33	30	1.15	1.45
Complaints	2	0.40	8	0.38	10	0.38	0.27
Total	497	100.00	2 101	100.00	2 598	100.00	

Authors' own.

At a thematic level, it was possible to identify a high variety of topics in the tweet sample. Besides the more expected aspects that connect local heritage and tourism in Lisbon and Alfama, such as fado, trams, landscape and traditional architecture, other topics emerged, relating to new patterns of urban tourism consumption such as contemporary cultural creation, including urban art (present in 5% of the thematic comments in tweets attributed to tourists), events such as the WebSummit (4%) and art, galleries and exhibitions (3%). More general aspects, such as those that showcase an historic urban landscape (HUL)³, the architecture, landscape or the atmosphere, appeared in almost one-third of the thematic comments (29.20%).

3. Historic urban landscape (HUL) is a new heritage approach focusing on the quality of human environment promoted by the UNESCO World Heritage Centre (UNESCO, 2012).

The differences between tourists and residents existed in relation to variables clearly associated with each group of users. For example, tourists were often associated with the tags “trips and visits” and “WebSummit” (an event that attracts a high number of foreign visitors to Lisbon), whereas residents were often the source of the messages tagged as “publicity” for events or businesses and “university”. However, a similar proportion of tourists and residents posted comments on “architecture and landscape”, its “atmosphere”, “Fado”, “restaurants and bars” and the “Santos Populares festivities”.

A linear discriminant analysis (LDA) revealed identical results: the most significant tags (positive coefficient) for tourists were those relating to the “WebSummit” and comments about “trips and visits”. On the other hand, the most significant tags for residents were “university” and “publicity”.

An analysis of location tags (Table 4) indicated that “Alfama” was the most frequent location (toponym in this case) for tourists and residents alike, occurring in 43% of tweets, followed by “Lisbon”, accounting for 14% (and 17% among tourists). References to other places were statistically less relevant. Therefore, Alfama appears as a specific spatial reference for tourists instead of blurring within the general reference of Lisbon. This situation is consistent with its high and growing tourist use but mainly reinforces the idea of the neighbourhood as a coherent whole.

Table 4. Alfama. Tweet content (location tag)

	Resident		Tourist		Total	
	Total	%	Total	%	Total	%
Portugal	1	0.09	27	1.03	28	0.75
Alfama	489	44.29	1109	42.14	1598	42.77
Lisboa	71	6.43	449	17.06	520	13.92
Other places	225	20,39	682	25,92	907	24,27
s.d.	318	28.80	365	13.87	683	18.28
Total	1 104	100.00	2 632	100.00	3 736	100.00

Authors' own.

VI. DISCUSSION AND CONCLUSIONS

The investigation demonstrated a method of exploring new data sources such as Twitter to advance knowledge on the activity patterns of tourist use at the destination places. These studies can provide valuable information for urban tourism planning and management. In this exploratory study, we focused on the potential of the social network, Twitter, as a data source to analyse tourist use of urban spaces. The case study allowed to propose a portrait for the city of Lisbon and its neighbourhood Alfama. The analyses revealed that Alfama concentrated (in the 2 years of harvested data) the highest tourist activity associated with Twitter users in Lisbon. The findings are consistent with the scenarios described while revisiting primary and secondary sources, which we presented in the introduction and places Lisbon and Alfama in a

network of world cities, together with their historic neighbourhoods, facing the impacts of mass tourism and the pressure of touristification.

The case of Lisbon is also consistent when we look at other big data analysis which retrieved geotagged photos (Encalada et al., 2017). All this tends to confirm the idea that the city undergoes an accelerated process of transformation triggered by the increasing flows of visitors. The multiple studies on the topic focusing on Lisbon's case are consistent at emphasizing the dangers of pursuing a path that tends to prioritise tourism over the accountability towards the needed for debate, considering alternative policies and change management (Malet Calvo, Gago & Cocola-Gant, 2018; Gago, 2018; Mendes, 2017).

The method employed in this study enabled us to identify areas in the city (almost in real-time) with the highest concentrations of resident and tourist Twitter users and classify these areas according to a greater or lesser influx of each type of user.

The analysis focused on tweets' content revealed traces of qualitative changes that affect tourist behaviour (new forms of urban tourism consumption). The Fado, the picturesque atmosphere, an historic urban landscape, but also urban art, new events, the proliferation of restaurants and comments on them indicate the activation of resources that serve as tourist attractions in a neighbourhood whose predominant use was, just a few years ago, housing. Thus, the case of Alfama presents particular aspects, as the community bounded to this neighbourhood sees it as their own territory, imprinting in the narrow streets their own personal character and letting their private lives move outside walls and doors. The Santos Populares' celebrations are an extreme case of such a relation with space.

Although the pressure cause by overtourism was not a recurrent theme in the sample of tweets analysed, comments began to appear expressing impatience and frustration ("Lisbon is still heaving with tourists", "Alfama in the morning: No tourists in sight?"). What should also be underpinned is that both tourists and residents alike made such comments. We can interpret that as a process of recentralisation occurring in these spaces; far from being abandoned by the local population as leisure spaces, they have their leisure and recreational functions reinforced as residents and tourists alike converge in increasingly overcrowded spaces as shown in the literature (Wilson & Tallon, 2012; Ashworth & Page, 2011; Hiernaux & González, 2014). It also reveals that residents adopt a "tourist view" of these central spaces in the process of touristification.

In addition, the location tag analysis proved that Alfama has a tourist identity since it was the most frequent spatial reference for tourists tweeting from there proving a scenario of touristification. However, we are aware that the temporal scope of the data sample would need to be extended in order to reach any definitive conclusions about these patterns of intensification in tourist use of the Alfama based on the increased presence of tourist Twitter users.

In conclusion, the study has demonstrated the suitability of the method used and some of the possible applications to urban spaces in other cities. The main advantages of this analysis are 1) the spatio-temporal disaggregation of data which allows identifying the daily activity patterns of users; 2) classifying Twitter users into residents and tourists and refining the definition of tourists by adding the category of national tourists in Lisbon while comparing their

spatial and temporal patterns (residents/tourists); 3) classifying areas in the city according to their tourist/residential specialisation; 4) determine patterns of urban tourism consumption through content analysis of Twitter messages. The use of extended time-frames would render more accurate simulations of processes of intensification in tourist and recreational use of the spaces considered, by applying similar mechanisms to analyse the content in association with the users' profiles and their role as non-residents.

This study takes a step forward by presenting patterns of tourists' use of space and conveyed narratives (content analysis) associated with the specific study case. The conclusions can raise decision makers' awareness, leading to a more informed planning and management, namely on what concerns the regulation and management of tourism resources, services and activities. In a context of sustained growth in tourism flows, it means not only an opportunity in terms of increased direct income, but also the foreseeable negative impacts should be tackled, to prevent scenarios where the quality of life of residents could be threatened and generating resistance movements in result.

The limitations applicable to this research display several challenges for future research. First, it is necessary to assume a population bias associated with the level of penetration of Twitter. The service attracts a specific profile of users (and therefore of tourists) representing a very limited age range (e.g. millennials, Generation Z) and high engagement with technology (e.g. heavy users of social media sites and digital devices). Second, there is a geographical bias associated with the different levels of penetration of this application within countries. Third, the data downloaded in real-time from Twitter's API only accounts for approximately 1% of the total tweets posted in the study period.

In the authors' opinion, social networks' data and particularly Twitter data can be used as a complementary source to official statistics and other methods of data collecting such as fieldwork. Despite the limitations represented by the sample's size, if the downloading process is performed for long periods of time, large data samples can be retrieved. Furthermore, the possibility of performing spatio-temporal disaggregation of geotagged information, combined with the fact that data can be captured in real-time, provides updated and disaggregated information on tourism flows, that can hardly be achieved with official statistics, and this constitutes a compelling advantage.

FUNDING DETAILS

DynMobility

Análisis dinámico de los patrones de movilidad a partir del Big Data.

Código: RTI2018-098402-B-I00.

Proyecto coordinado por: Javier Gutiérrez Puebla (tGIS) y Juan Carlos Martín (ITyT - ULPGC).

Entidad financiadora: Ministerio de Ciencia, Innovación y Universidades (Plan Nacional).

Periodo de realización: 2019-2021.

Partners: ITyT (ULPGC) y G-TEC Grasia (UCM).

PAITUR-Ciudad

Transformaciones del Paisaje Urbano Histórico inducidas por el turismo: contradicciones y controversias, gobierno y gobernanza Local.

Código: CSO2016-75470-R.

Proyecto coordinado por: Manuel de la Calle Vaquero (Turismo, Patrimonio y Desarrollo) y María Velasco González (Espacio y poder).

Entidad financiadora: Ministerio de Ciencia, Innovación y Universidades (Plan Nacional).

Periodo de realización: 2017-2021.

Beca de investigación predoctoral

PhD Scholarship.

Código: SFRH/BD/121371/2016.

Organismo: FCT – Fundação para a Ciência e Tecnologia (Portugal).

REFERENCES

- Arias-Sans, A., & Quaglieri-Domínguez, A. (2016). Unravelling Airbnb: urban perspectives from Barcelona. In A. P. Russo & G. Richards (Eds.), *Reinventing the local in tourism. Travel Communities and Peer-Produced Place Experiences* (pp. 209-228). London: Channel View Publications.
- Ashworth, G., & Page, S. J. (2011). Urban tourism research: Recent progress and current paradoxes. *Tourism Management*, 32(1), 1-15.
- Bae, S. H., & Yun, H. J. (2017). Spatiotemporal Distribution of Visitors' Geotagged Landscape Photos in Rural Areas. *Tourism Planning and Development*, 14(2), 167-180. <https://doi.org/10.1080/21568316.2016.1204356>
- Baggio, R. (2016). Big Data, Business Intelligence and Tourism: a brief analysis of the literature. *Big Data & Business Intelligence in the Travel & Tourism Domain*, 1-7. <https://doi.org/10.1016/j.tourman.2015.10.008>
- Barata-Salgeiro, T. (2017). Alojamentos turísticos em Lisboa. *Scripta Nova: Revista Electrónica de Geografía Y Ciencias Sociales*, XXI(578), 1-43.
- Bassolas, A., Lenormand, M., Tugores, A., Gonçalves, B., & Ramasco, J. J. (2016). Touristic site attractiveness seen through Twitter. *EPJ Data Science*, 5(1). <https://doi.org/10.1140/epjds/s13688-016-0073-5>
- Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, 68, 101-115. <https://doi.org/10.1016/j.tourman.2018.02.020>
- Benis, K. (2011). *Vielas de Alfama. Entre Revitalização e Gentrificação. Impactos da «gentrificação» sobre a apropriação do espaço público*. Universidade Técnica de Lisboa.
- Birenboim, A., Reinau, K. H., Shoval, N., & Harder, H. (2015). High-Resolution Measurement and Analysis of Visitor Experiences in Time and Space: The Case of Aalborg Zoo in Denmark. *Professional Geographer*, 67(4), 620-629. <https://doi.org/10.1080/00330124.2015.1032874>
- Brogueira, G., Batista, F., & Carvalho, J. P. (2016). Using geolocated tweets for characterization of Twitter in Portugal and the Portuguese administrative regions. *Social Network Analysis and Mining*, 6(1). <https://doi.org/10.1007/s13278-016-0347-8>
- Calle Vaquero, M. de la. (2002). *La Ciudad Histórica como Destino Turístico*. Barcelona: Editorial Ariel.

- Canalis, X. (2018). Las tendencias mundiales del turismo que veremos en la ITB. *Hosteltur*. Retrieved from https://www.hosteltur.com/126641_tendencias-mundiales-turismo-veremos-itb.html
- Chua, A., Servillo, L., Marcheggiani, E., & Moere, A. Vande. (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tourism Management*, 57, 295-310. <https://doi.org/10.1016/j.tourman.2016.06.013>
- Clerval, A., Colomb, C., & Crieckingen, M. van. (2011). La gentrification des métropoles européennes. In *Données urbaines* (pp. 151-165). Economica Anthropos.
- Cócola Gant, A. (2015). Tourism and commercial gentrification. In *RC21 International Conference on "The Ideal City: between myth and reality. Representations, policies, contradictions and challenges for tomorrow's urban life* (pp. 27-29). Retrieved from <http://www.rc21.org/en/conferences/urbino2015>
- Colomb, C. (2012). A city for whom? Urban tourism, neighbourhood conflicts and planning policy in Berlin and Barcelona. In *Second ISA Forum of Sociology (08/01-04)*. Buenos Aires.
- Colomb, C., & Novy, J. (2016). *Protest and resistance in the tourist city*. Abingdon, Oxon & New York, NY: Routledge.
- Comito, C., Falcone, D., & Talia, D. (2016). Mining human mobility patterns from social geotagged data. *Pervasive and Mobile Computing*, 33(September 2017), 91-107. <https://doi.org/10.1016/j.pmcj.2016.06.005>
- Cordeiro, G., & Costa, A. F. (2003). Territorio, identidad y sociedades de barrio. In J. I. Homobono & J. A. Rubio-Ardanaz (Eds.), *Las culturas de la ciudad* (pp. 763-785). Donostia: Eusko Ikaskuntza.
- Crandall, D. J., Backstrom, L., Huttenlocher, D., & Kleinberg, J. (2009). Mapping the world's photos, 761. <https://doi.org/10.1145/1526709.1526812>
- Davidson, M., & Lees, L. (2005). New build "gentrification" and London's riverside renaissance. *Environment and Planning A*, 37(7), 1165-1190.
- Donaire, J. A., Camprubí, R., & Galí, N. (2014). Tourist clusters from Flickr travel photography. *Tourism Management Perspectives*, 11, 26-33. <https://doi.org/10.1016/j.tmp.2014.02.003>
- Encalada, L., Boavida-Portugal, I., Ferreira, C. C., & Rocha, J. (2017). Identifying tourist places of interest based on digital imprints: Towards a sustainable smart City. *Sustainability*, 9(12). <https://doi.org/10.3390/su9122317>
- Feick, R., & Robertson, C. (2015). A multi-scale approach to exploring urban places in geotagged photographs. *Computers, Environment and Urban Systems*, 53, 96-109. <https://doi.org/10.1016/j.compenvurbsys.2013.11.006>
- Firmino da Costa, A. (2008). *Sociedade de Bairro. Dinâmicas sociais da identidade cultural*. Oeiras: Celta Editora.
- Florida, R. (2002). *The rise of the creative class: and how it's transforming work, leisure, community and everyday life*. New York: Basic Books.
- Fontes, A. C. (2021) *As pequenas escalas da rua. Morfologia, apropriação e significados do beco em Lisboa*. Tese de Doutoramento em Estudos Urbanos, ISCTE-IUL and FCSH-UNL.
- Füller, H., & Michel, B. (2014). 'Stop being a tourist!' New dynamics of urban tourism in Berlin-Kreuzberg. *International Journal of Urban and Regional Research*, 38(4), 1304-1318.

- Gago, A. (2018). *O aluguer de curta duração e a gentrificação turística em Alfama, Lisboa*. Universidade de Lisboa.
- Garay, L., & Morales Pérez, S. (2017). Understanding the creation of destination images through a festival's Twitter conversation. *International Journal of Event and Festival Management*, 8(1), 39-54. <https://doi.org/10.1108/IJEFM-04-2016-0030>
- García-Palomares, J. C., Gutiérrez, J., & Mínguez, C. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Applied Geography*, 63, 408-417. <https://doi.org/10.1016/j.apgeog.2015.08.002>
- Girardin Fabien, B. J., Calabrese, F., Fiore, F. D., & Ratti, C. (2008). Digital Footprinting: Uncovering Tourists with User-Generated Content. *IEEE Pervasive Computing*, 36-43.
- Glass, R. (1964). Introduction: aspects of change. In Centre for Urban Studies (Ed.), *Aspects of Change* (pp. xiii-xlii). London: MacGibbon and Kee.
- Gotham, K. F. (2005). Tourism gentrification: The case of new Orleans' Vieux Carré (French Quarter). *Urban Studies*, 42(7), 1099-1121.
- Hartman, C., Keating, D., & LeGates, R. (1982). *Displacement: How to Fight It*. Berkeley, CA: National Housing Law Project.
- Hauthal, E., & Burghardt, D. (2014). Mapping Space-Related Emotions out of User-Generated Photo Metadata Considering Grammatical Issues, *O(0)*, 1-13. <https://doi.org/10.1179/1743277414Y.0000000094>
- Hawelka, B., Sitko, I., Beinart, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260-271. <https://doi.org/10.1080/15230406.2014.890072>
- Hedrick-Wong, Y., & Choong, D. (2016). Global Destination Cities Index. Mastercard.
- Henriques, E. B. (1996). *Lisboa turística - entre o imaginário e a cidade*. Lisboa: Edições Colibri.
- Hiernaux, D., & González, I. (2014). Turismo y gentrificación: pistas teóricas sobre una articulación. *Revista de Geografía Norte Grande*, 58, 55-70.
- Hu, Y., Gao, S., Janowicz, K., Yu, B., Li, W., & Prasad, S. (2015). Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems*, 54, 240-254. <https://doi.org/10.1016/j.compenvurbsys.2015.09.001>
- Judd, D. R. (2003). Visitors and the spatial ecology of the city. In J. Hoffman, S. Fainstein, & D. R. Judd (Eds.), *Cities and visitors: Regulating people, markets, and city space*. Oxford: Blackwell.
- Kádár, B. (2014). Measuring tourist activities in cities using geotagged photography. *Tourism Geographies*, 16(1), 88-104. <https://doi.org/10.1080/14616688.2013.868029>
- Kesar, O., Deželjin, R., & Bienenfeld, M. (2011). Tourism gentrification in the city of Zagreb: Time for a debate? *Interdisciplinary Management Research*, 657-668.
- Kim, H., & Stepchenkova, S. (2015). Effect of tourist photographs on attitudes towards destination: Manifest and latent content. *Tourism Management*, 49, 29-41. <https://doi.org/10.1016/j.tourman.2015.02.004>
- Lees, L. (2008). Gentrification and social mixing: towards an urban renaissance? *Urban Studies*, 45(12), 2449-2470.
- Lera I., Pérez T., Guerrero C., Eguíluz V., J. C. (2017). Analysing human mobility patterns of hiking activities through complex network theory. *PLoS ONE*, 12(5).

- Lestegás, I., Seixas, J., & Lois-González, R.-C. (2019). Commodifying Lisbon: A Study on the Spatial Concentration of Short-Term Rentals. *Social Sciences*, 8(2), 33. <https://doi.org/10.3390/socsci8020033>
- Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, 68, 301-323. <https://doi.org/10.1016/j.tourman.2018.03.009>
- Liu, F., Zhu, X., Li, J., Sun, J., & Huang, Q. (2019). Progress of gentrification research in China: A bibliometric review. *Sustainability (Switzerland)*, 11(2). <https://doi.org/10.3390/su11020367>
- Lloyd, R., & Clark, T. N. (2001). The city as an entertainment machine. In K. F. Gotham (Ed.), *Critical perspectives on urban redevelopment* (Vol. 6, pp. 357-378). New York: Elsevier.
- Mahmud, J., Nichols, J., & Drews, C. (2014). Home Location Identification of Twitter Users. *ACM Transactions on Intelligent Systems and Technology*, 5(3), 1-21. <https://doi.org/10.1145/2528548>
- Maitland, R., & Newman, P. (2009). *World Tourism Cities: developing Tourism off the beaten track*. New York: Routledge.
- Malet Calvo, F., Gago, A., & Cocola-Gant, A. (2018). Turismo, negocio inmobiliario y movimientos de resistencia en Lisboa, Portugal. In C Milano & J. A. Mansilla (Eds.), *Ciudad de Vacaciones. Conflictos urbanos en espacios turísticos* (Observator). Barcelona: Pol.len Edicions.
- Manca, M., Boratto, L., Morell Roman, V., Martori i Gallissà, O., & Kaltenbrunner, A. (2017). Using social media to characterize urban mobility patterns: State-of-the-art survey and case-study. *Online Social Networks and Media*, 1(December), 56-69. <https://doi.org/10.1016/j.osnem.2017.04.002>
- Marine-Roig, E. (2017). Measuring destination image through travel reviews in search engines. *Sustainability (Switzerland)*, 9(8). <https://doi.org/10.3390/su9081425>
- Marine-Roig, E., & Anton Clavé, S. (2015). Tourism analytics with massive user-generated content: A case study of Barcelona. *Journal of Destination Marketing and Management*, 4(3), 162-172. <https://doi.org/10.1016/j.jdmm.2015.06.004>
- Mendes, L. (2017). Gentrificação turística em Lisboa: neoliberalismo, financeirização e urbanismo austeritário em tempos de pós-crise capitalista 2008-2009. *Cadernos Metrópole*, 19(39), 479-512. <https://doi.org/10.7476/9788523013400.0007.01>
- Mendes, L. (2020). Bye Bye Lisbon: Tourism Gentrification Impacts on Lisbon's Inner-City Housing Market. In C. Ribeiro de Almeida, A. Quintano, M. Simancas, R. Huete, & Z. Breda (Eds.), *Handbook of Research on the Impacts, Challenges, and Policy Responses to Overtourism* (pp. 136-155). IGI Global Publisher. <https://doi.org/10.4018/978-1-7998-2224-0>
- Messe Berlin. (2018). ITB Berlin and IPK International : International City Trips - A Success Story. *ITB Berlin*. Retrieved from https://www.itb-berlin.de/Presse/Pressemitteilungen/PDF_56725.jsp
- Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A Big Data Analytics Method for Tourist Behaviour Analysis. *Information and Management*, 54(6), 771-785. <https://doi.org/10.1016/j.im.2016.11.011>
- Milano, Claudio, Novelli, M., & Cheer, J. M. (2019). Overtourism and Tourismphobia: A Journey Through Four Decades of Tourism Development, Planning and Local Concerns. *Tourism Planning & Development*, 16(4), 353-357.

- Morell, M. (2009). Fent barri: heritage tourism policy and neighbourhood scaling in Ciutat de Mallorca. *Etnogràfica*, 13(2), 343-372.
- Murthy, D. (2013). *Twitter: Social communication in the twitter age*. John Wiley & Sons.
- Okazaki, S., Andreu, L., & Campo, S. (2017). Knowledge Sharing Among Tourists via Social Media: A Comparison Between Facebook and TripAdvisor. *International Journal of Tourism Research*, 19(1), 107-119. <https://doi.org/10.1002/jtr.2090>
- Pantano, E., Priporas, C. V., & Stylos, N. (2017). 'You will like it!' using open data to predict tourists' response to a tourist attraction. *Tourism Management*, 60, 430-438. <https://doi.org/10.1016/j.tourman.2016.12.020>
- Pérez Guilarte, Y. & Lois González, R.C. (2018) Sustainability and visitor management in tourist historic cities: the case of Santiago de Compostela, Spain, *Journal of Heritage Tourism*, 13:6, 489-505, <https://doi.org/10.1080/1743873X.2018.1435665>.
- Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management*, 55, 16-24. <https://doi.org/10.1016/j.ijhm.2016.02.001>
- Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202-212. <https://doi.org/10.1016/j.tourman.2016.06.006>
- Russo, A. P., & Quaglieri-Domínguez, A. (2014). La lógica espacial del intercambio de casas: una aproximación a las nuevas geografías de lo cotidiano en el turismo contemporáneo. *Scripta Nova: Revista Electrónica de Geografía Y Ciencias Sociales*, 18(483).
- Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2018). Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management*, 66, 13-25. <https://doi.org/10.1016/j.tourman.2017.11.001>
- Sequera, J., & Janoshka, M. (2013). Gentrification dispositifs in the historic centre of Madrid: a reconsideration of urban governmentality and state-led urban reconfiguration. In L. Lees, H. Bang Shin, & E. López-Morales (Eds.), *Urban gentrifications: uneven development and displacement*. London: Routledge.
- Shao, H., Zhang, Y., & Li, W. (2017). Extraction and analysis of city's tourism districts based on social media data. *Computers, Environment and Urban Systems*, 65, 66-78. <https://doi.org/10.1016/j.compenvurbsys.2017.04.010>
- Shi, B., Zhao, J., & Chen, P. J. (2017). Exploring urban tourism crowding in Shanghai via crowdsourcing geospatial data. *Current Issues in Tourism*, 20(11), 1186-1209. <https://doi.org/10.1080/13683500.2016.1224820>
- Shoval, N. (2018). Sensing tourists: geoinformatics and the future of tourism geography research. *Tourism Geographies*, 20(5), 910-912. <https://doi.org/10.1080/14616688.2018.1437768>
- Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: the first decade. *Tourism Geographies*, 18(5), 587-606. <https://doi.org/10.1080/14616688.2016.1214977>
- Shoval, N., Schvimer, Y., & Tamir, M. (2018). Real-Time Measurement of Tourists' Objective and Subjective Emotions in Time and Space. *Journal of Travel Research*, 57(1), 3-16. <https://doi.org/10.1177/0047287517691155>
- Slater, T. (2009). Missing Marcuse: on gentrification and displacement. *City*, 13(2), 292-311.

- Sotiriadis, M. D., & van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: The use of twitter by tourists. *Electronic Commerce Research*, 13(1), 103-124. <https://doi.org/10.1007/s10660-013-9108-1>
- UNESCO. (2012). Records of the General Conference, 36th session. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO). Retrieved from http://portal.unesco.org/en/ev.php-URL_ID=48857&URL_DO=DO_TOPIC&URL_SECTION=201.html
- UNWTO. (2018). 'Overtourism'? Understanding and Managing Urban Tourism Growth beyond Perceptions Volume 2: Case Studies. World Tourism Organization. <https://doi.org/10.18111/9789284420643>
- Urry, J. (1990). *The tourist gaze. Leisure and travel in contemporary societies*. London: Sage Publications.
- van der Zee, E., Bertocchi, D., & Vanneste, D. (2018). Distribution of tourists within urban heritage destinations: a hot spot/cold spot analysis of TripAdvisor data as support for destination management. *Current Issues in Tourism*, 0(0), 1-22. <https://doi.org/10.1080/13683500.2018.1491955>
- Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management*, 46, 222-232. <https://doi.org/10.1016/j.tourman.2014.07.003>
- Williams, N. L., Inversini, A., Ferdinand, N., & Buhalis, D. (2017). Destination eWOM: A macro and meso network approach? *Annals of Tourism Research*, 64, 87-101. <https://doi.org/10.1016/j.annals.2017.02.007>
- Wilson, J., & Tallon, A. (2012). Geographies of gentrification and tourism. In J. Wilson (Ed.), *The Routledge Handbook of Tourism Geographies* (pp. 103-112). London: Routledge.
- Zakrisson, I., & Zillinger, M. (2012). Emotions in motion: Tourist experiences in time and space. *Current Issues in Tourism*, 15(6), 505-523. <https://doi.org/10.1080/13683500.2011.615391>
- Zhang, J. (2018). Big data and tourism geographies – an emerging paradigm for future study? *Tourism Geographies*, 20(5), 899-904. <https://doi.org/10.1080/14616688.2018.1519719>

CONTRIBUCIONES

Authors contribution

Claudia Yubero, Ana M. Condeço-Melhorado, María García-Hernández and Ana Catarina Fontes have conceived and structured the research, reviewed the literature, designed the methodology, collected and analysed the data. All four authors have written and critically reviewed the paper, also in its final version.