DISCOVERING UGC COMMUNITIES TO DRIVE MARKETING STRATEGIES: LEVERAGING DATA VISUALIZATION

Jose Ramon Saura[1]
Ana Reyes-Menendez[2]
Pedro Palos-Sanchez[3]
Ferrão Filipe[4]

ABSTRACT

The digital tourism ecosystem is changing driven by the massive use of new technologies and the increase of data generated by travelers during their trips. In this digital landscape, businesses in the tourism sector are adapting their strategies to take better advantage of new knowledge that can be extracted from the Internet, specifically from social networks. The objective of this research is to define the content strategy in social networks that businesses in the digital tourism sector should follow and to highlight the importance of new data visualization techniques for Marketing and Marketing analytics. To this end, tourist communities of Twitter users have been detected by analyzing the User Generated Content (UGC) and applying algorithms for data visualization on a sample of n = 10.00 tweets from the interactions between the UGC on Twitter and the 25 Top Hotels in the World as designated by TripAdvisor's Travelers' Choice Awards. A total of n = 3.158 tweets were analyzed, focusing specifically on comments that had hashtags and interactions with these hotels, with the aim of detecting communities according to the type of content shared, in order to measure the communities' impact and influence in the digital tourism sector. The results of this research study identify the main topics related to the tourism sector and the most active communities according to their impact, which allows CEOs and managers of tourism companies to refine their marketing strategies for the digital tourism industry.

Keywords: UGC Communities, Marketing Strategies, Data Visualization, Marketing Analytics.

JEL Classification: M2, M3

1. INTRODUCTION

It is expected that 50 times more data will be generated in 2020 than previously in 2011. It is also expected that there will be 75 times more information sources compared to the present, and that each inhabitant on planet Earth will have 6.58 devices connected to the Internet (Kim, 2016).

Companies have realized that the large amounts of data that make up Big Data datasets allow them to better know their customers, make better decisions, and improve user experience (Jovicic, 2019). That is why companies from different sectors, especially those focused on client experience in the tourism sector, are using Big Data and Business Intelligence (BI) to

[1] Department of Business Economics, Faculty of Social Sciences and Law, Rey Juan Carlos University, Madrid, Spain (joseramon.saura@urjc.es)
[2] Department of Business Economics, Faculty of Social Sciences and Law, Rey Juan Carlos University, Madrid, Spain (ana.reyes@urjc.es)
[3] Department of Business Administration and Marketing, University of Sevilla, Seville, Spain (ppalos@us.es)
[4] Vice-Rector, Universidade Portucalense, Porto, Portugal (ferraof@upt.pt)

improve decision making (Lunar & Jacobsen, 2014). The problem with these companies is that the tools they use to extract information from the data are too complex, require advanced technical knowledge, or require a lengthy reporting period (Murray, 2015).

Clients expect that the information a company has about them will be used to provide them better service; for that reason, the Real Time Big Data (RTB) is one of the main trends within the management of large amounts of data (Palos-Sanchez et al., 2019). Clients share information with the company and in return expect companies to use that information to offer them a better or more personalized experience, or to provide a solution to their problems in real time (Liu et al., 2013). This process is especially important in sectors such as tourism, where the client is at the center of the strategy (Reyes-Menendez et al., 2018) and the key lies in experience.

Another main trend within the management of Big Data is the management of data produced in mobile environments, that is, through the connected devices that users have. These devices are mainly smartphones that users keep on hand to carry out their daily activities and produce large amounts of data. As for the Mobile Big Data (MBD), it is interesting to highlight the use of Social Networks by users (Hays et al., 2013).

Consumers share valuable information online, called User Generated Content (UGC), for tourism companies on social platforms. UGC is defined as "the content generated by users in social networks and digital platforms. Such content includes comments, opinions, expressions, and interactions between users and brands, or any other type of content shared publicly on the Internet that seeks to generate engagement between different profiles" (Saura & Bennett, 2019: 4). Companies should use that UGC to improve services offered in order to improve decision-making processes, streamline strategies, and obtain insights as stated by Saura and Bennett (2019). Insights are key to the consumers or users that are obtained when analyzing large amounts of data (Saura et al., 2019) in the Tourism industry.

In this context, data visualization tools that allow extraction of key information about the UGC while not requiring advanced tool programming are allowing companies to find the balance between the amount of data used and the capacity to obtain agile information key to the company (Brizirgianni & Dionysopoulou, 2013; Reyes-Menendez et al., 2019). Therefore, the main objective of this research is to present different data visualization algorithms based on UGC applicable to the Tourism sector with the exploratory purpose of understanding the results of the UGC analysis in Social Media applied to Tourism. To cover this objective, we propose 2 research questions; RQ1: Is it possible to apply data visualization algorithms in Twitter UGC to obtain insights that can be used by CEOs and Executives of tourism companies? and RQ2: Can we analyze the user communities that interact on Twitter around the tourist-type UGC and discover useful knowledge for tourism businesses?

To understand the importance of these visualization techniques and the analysis of UGC communities for the tourism sector, we collected a total of n = 10,000 tweets from interactions between the UGC of users on the social network Twitter and the 25 Top Hotels in the World as designated by TripAdvisor's Travelers' Choice Award. Of the total of n = 10,000 tweets, a total of n = 3,158 tweets were analyzed, focusing specifically on comments that had hashtags, with the objective of detecting communities according to the type of content shared.

The remainder or this manuscript is structured as follow: firstly, we present Introduction, next the Literature Review is presented divided into two subparts, data visualization for marketing and tourism and social media data visualization. Third, the methodological process is presented in which the algorithms used and known as Graph Distance Report (GDR), Eigenvector Centrality Report (ECR), HITS metric and Authority Distribution

(AD) are introduced. Finally, the conclusions in which practical and theoretical implications of research are highlighted.

2. LITERATURE REVIEW

2.1 Data visualization for Marketing

Shaw et al. (2001) perform a methodology based on data collection and the use of knowledge management techniques in order to manage marketing knowledge and make correct decisions for the company's purposes (Xiang et al., 2017). Likewise, Few and Edge (2007) carry out a study on the present, past, and future of data visualization in order to understand how to properly use the information provided to the company, both to make sense of the data and to enable action after the analysis of data dashboards.

Likewise, Maimon and Rokach (2005) develop a manual on Data Mining (DM) and Knowledge Discovery in Databases (KDD), organizing the concepts, standards, and most current applications on the DM for correct understanding of these marketing concepts by the company.

Following this line of research, Reddy et al. (2019) develop a survey on Business Intelligence (BI) tools for three different areas: marketing services, transport services, and financial services. Vassakis et al. (2019) develop a methodology for obtaining and analyzing the visualization of data extracted from social networks. This study focuses on the tourism sector by analyzing customer experiences.

Zarco et al. (2019) conduct a study focused on the collection of data in the social network Twitter in traditional companies, seeking the impact and positioning of each company in the social network via methods and analysis techniques that show the relevance of techniques and methods of data analysis for marketing and tourism sectors, among others.

2.2 Social Media Data Visualization for Marketing and Tourism

The research of Jimenez-Marquez et al. (2019) present an approach to analyzing content created by users in social networks to identify the value of services and products of companies, with the aim of improving their milestones in the era of big data. Also, Börner et al. (2019) classify different definitions and assessments about Data Visualization Literacy (DVL) in order to teach their application and use. They focus the study with practical exercises and examples across different sectors and industries.

Saura et al. (2018) develops a study focused on the analysis of UGC on the Twitter social network identifying comments between hotels and users, using as an international hotel ranking show. In the investigation, data analysis techniques are used showing importance to data visualization and analysis techniques (Leung et al., 2013).

Kucher (2019) performs an analysis on the data extracted by the "Social Media" using visualization techniques based on case studies and visual representations of methods for obtaining data. In this way, and also following these methodological approaches based on data, Confente et al. (2019) elaborate an analysis of corporate reputation in order to identify the differences about UGC in social networks. They then study the results, evaluating clients' opinions by performing data visualization techniques.

Likewise, Reyes-Menendez et al. (2019) develops an investigation focused on the analysis of the #WorldEnvironmentDay hagstag in which conclusions are obtained regarding the decision-making for executives linked to tourism providing information that concerns the geographical areas in which the hotel carry out their activities. Also, Saura et al. (2018) use the UGC of comments extracted from the social network TripAdvisor to obtain key indicators related to the tourism sector analyzing the comments of the users who publish during their trips on social networks.

3. METHODOLOGY

The main objective of this research is exploratory and based on knowledge discovery, not hypothesis testing and not trying to control variables, but to discover them (Corbin & Strauss, 2015). The methodology is based on data visualization techniques and knowledge extraction from databases focused on UGC analysis following the indications presented by Saura and Bennett (2019) with the purpose of helping CEOs and executives to make better marketing decisions based in data analysis.

For data collection, we connected to the public Twitter API from September 3 to September 10, 2018 with a total of n = 10,000 tweets; the downloaded tweets reflect UGC published on Twitter and represent interaction with hotels included in the Top 25 Hotels in the World as designated by TripAdvisor Travelers' Choice Awards. These days they were selected randomly estimating that it did not coincide with any world tourism event with repercussion on digital or social media channels that could have affected the sample (Jia, 2018). After cleaning the dataset, a total of n = 3,158 tweets were analyzed focusing specifically on comments that had hashtags with the objective of detecting communities according to the type of content shared.

For data visualization we have used the Gephi software, applying the following algorithms and experiments to the dataset since all of them are Open Source type and can be used by researchers: Graph Distance Report, closeness centrality distribution and harmonic closeness centrality distribution after applying the algorithm presented by Brandes (2001); Eigenvector Centrality Report, to measure the eigenvector centrality distribution; HITS metric applying the Kleinberg, algorithm to measure the Hubs distribution and Authority Distribution; modularity report to measure the size distribution by applying the algorithm of Blondel et al. (2008) with the resolution of visualization of Lambiotte and Delvenne data (2009).

For the process of cleaning the UGC database, we followed the recommendations of Reyes-Menendez et al. (2019) and Saura and Bennett (2019). The exploratory analysis process is based on the analysis of the indicated databases classified in nodes. The nodes are small groups of information groups that link information related to different users in social networks, obtaining and being able to represent the bidirectionality of information and interaction between the users that make up a node (Saura & Bennett, 2019; Lio et al. 2019). Although they are dynamic databases, they can be analyzed visually representing the distances that exist between the main themes that compose them.

3.1 Graph Distance Report (GDR)

The GDR measures the distance between the nodes that compose the dataset. The aim of GDR is to identify similar patterns or linkages between node unions; it was developed by Brandes (2001).

The GDR visually identifies the connected nodes according to diameter. The diameter is the longest distance between the two nodes of the network, visualizing the distance in which the two nodes are most distant from a given dataset. In this case, the measures that can be identified are the distance Betweenness Centrality (BC) and Closeness Centrality (CC).

As indicated by Brandes (2001), BC represents the measure of the frequency with which a node appears in the shortest routes between the nodes of a network. Likewise, the average or DC indicator measures the average distance of a given node to all the others in the network. It is also possible to measure the distance from a specific node to the farthest node by applying filters based on Eccentricity (EC).

3.2 Eigenvector Centrality Report (ECR)

The ECR measures the importance of the node in a network based on node connections in order to identify its importance. Within graphics theory, the centrality of the eigenvector (also called eigenntrality) is a measure of the influence of a node on a network. Therefore, by applying ECR one can learn what the influence of a node is on the rest of the nodes of a given dataset (Mono & Tribe, 2017). This development allows for one to measure the similarity of contents or influences of these in the composition and grouping of other nodes.

When applying ECR, relative scores are assigned to all nodes in the network according to the concept that connections to high-score nodes contribute more to the score of the node in question than to connections equal to low-score nodes. A high own-vector score means that a node is connected to many nodes that have high scores so that the similarity of the influences can be identified. Likewise, it should be indicated that in neuroscience, it has been found that the centrality of a neuron's own vector, in a model neural network or a standard dataset, can be correlated with its trigger speed or the speed at which it grows.

3.3 HITS metric report

The HITS is a data analysis and visualization algorithm that measures the links that connect different web pages or nodes of a dataset. The HITS was developed by Kleinberg (1999) with the objective of determining what the values of a page or node are in terms of metrics or key indicators related to authority, the estimation of the value of the content of the page, the value of the central nodes in the dataset, and the number of links that connect to other pages. Specifically, by applying the HITS algorithm, pages or nodes can be classified according to their authority by obtaining the HITS Authority distribution.

The Hubs distribution calculates globally two concrete scores in a differentiated way– on one hand, the centers and on the other, the authority of these. The authority score indicates the value of the page or node itself, and the distribution centers estimate the value of the outgoing links of the page or node, allowing to identify patterns as well as trends.

3.4 Modularity report

The MR was presented by Blondel (2008) and subsequently improved for visualization according to the type of resolution established by Lambiotte and Delvenne (2009). The main objective of the MR is to measure the decomposition of a network of nodes in modular communities, represent the communities based on the measurement of connections and interactions of these communities between the nodes.

Also, the MR indicates that if a dataset has a modularity score it indicates a sophisticated internal structure. This structure, often called community structure, describes how the network is divided into sub-networks that group nodes. It has been shown that these subnets (or communities) have significant meaning in the real world and in the analysis of social networks; thus, they can be used to analyze behavior or understand the contents around a community.

For the application of the algorithm, a random assignment can produce a better decomposition resulting from a higher modularity score, however, the randomization will increase the calculation time which may in turn allow identification of smaller communities centered on specific, or at least more specific, topics that do not have as much weight.

4. EXPLORATORY ANALYSIS OF RESULTS

4.1 Graph Distance Report

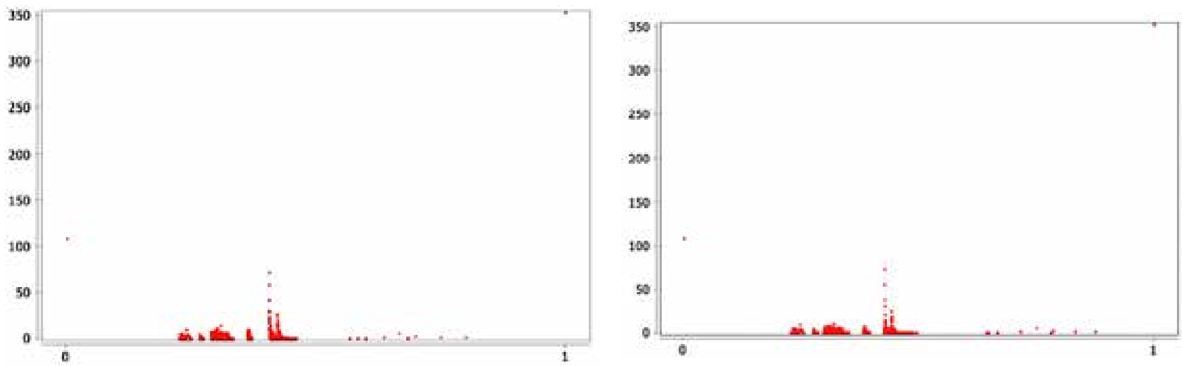
After applying the corresponding algorithm to GDR in (a), in order to be close to the subject of tourism on Twitter, topics related to travelers, vacations, photography, nature, trips, and exploration have been identified. This topics discovery allows us to understand the chatting in UGC around tourism in which comments and opinions are shared regarding the topics identified by exchanging these comments with the hotels.

Furthermore, if we look at the results of (b) shown by the HCCD, we can obtain a different order of identified topics such as travelers, photography, tours, vacations, trip, exploration, and nature, in which the subject matter of photography and photography becomes stronger. This fact shows that in the UGC collected sample, photography is considered a relevant issue when interacting with hotels and sharing opinions. Tours, as similar topics, are not directly centralized in tourism.

In this case, as indicated before, an application of this algorithm has been made with the Gephi software applying an interpretation of undirected parameters. In the results we have obtained a 6-point diameter, a Radius of 0, and an Average Path length of 2.77.

In Figure 1 (a) the Closeness Centrality Distribution is available, and in Figure 1 (b) the Harmonic Closeness Centrality Distribution is shown, in which the weight of the themes, identified with respect to the main theme of studio, tourism can be identified on the X axis.

Figure 1. Closeness Centrality Distribution and Harmonic Closeness Centrality Distribution



(a) (b)

Source: Own Elaboration

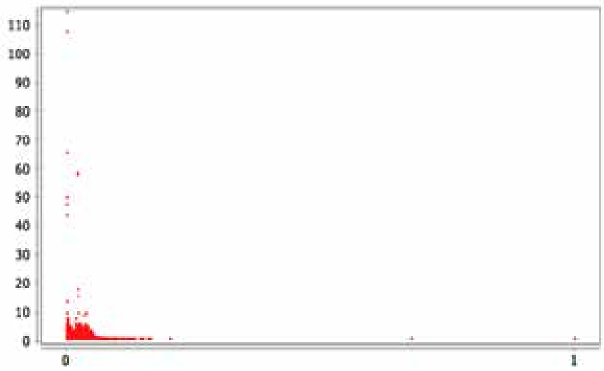
4.2 Eigenvector Centrality Report

The results show that the nodes that have more importance in terms of weight of interactions related to tourism are those related to instatravel, with a weight of 0.145, nature (0.165); tourist (0.0724); traveler (0.0844); frizemedia (0.0863); photography (0.1144); and vacation (0.1453).

It is interesting that in measuring the influence of the topics and nodes, we found that the appearance of the theme of #instatravel gains weight. This demonstrates that there is a high engagement in those users talking about and in which their tweets talk about tourism and contain images related to photographs taken during the trips and posted through hashtag, according to our sample.

To visualize this data, we applied network undirected interpretation parameters in which we have obtained a number of iterations of 1.111.001 and a sum change: 0.0. Figure 2 shows the result of ECD in the dataset showing on the X axis the total number of topics and their importance regarding the central point of the study, tourism, represented on the Y axis with 0 points.

Figure 2. Eigenvector Centrality Distribution



Source: Own Elaboration

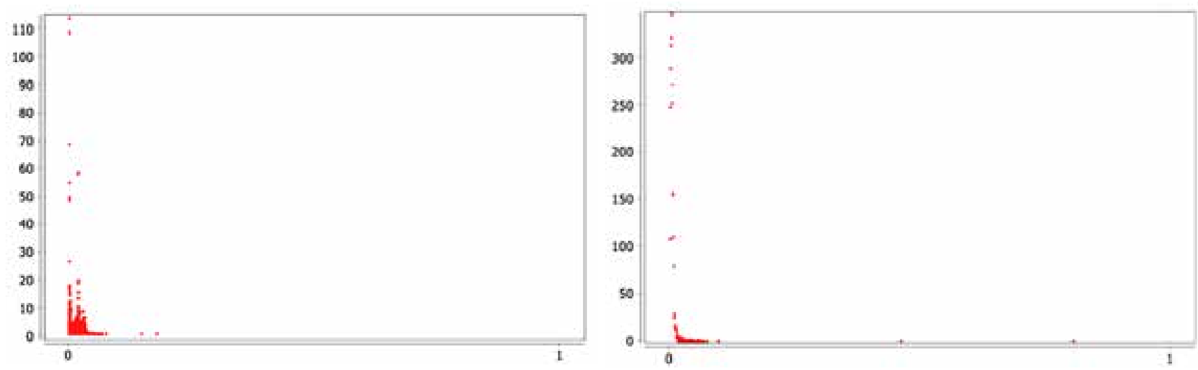
4.3 HITS metric report

In (a) the results related to the HITS distribution have identified the nodes with more links. The strongest results of HITS metric report is traveling, with an importance of (0.0848); trip (0.0764); tourist (0.0748); traveling (0.0704); holiday (0.0700); beach (0.0685); tour (0.068); and photography (0.0643). HITS distribution allows us to note that if the goal of a social media strategy for digital tourism is to get traffic to web pages born of UGC, the topics could contain content related to trip, tourist, and traveling to engage with user's, as the explorative results show.

Regarding the authority in social media (b) that is acquired when sharing content, the results have been the following: travelgram (0.0562); be (0.0541); adventure (0.0559); explore (0.0619); photography (0.0643); tour (0.0683); and beach (0.0685). So, companies that want to gain authority in social networks within the tourism ecosystem could use content to publish in their UGC strategies relative to similar topics to travelgram, adventure, and exploration.

In this way, in Figure 3 (a) the HITS distribution can be observed in (b) the HITS authority distribution.

Figure 3. HITS distribution (a) and Hits authority distribution (b)



(a) (b)

Source: Own Elaboration

4.4 Modularity report

Two results have been obtained to measure the number of UGC communities around digital tourism on Twitter; those are available in (a) and (b). We can consider that the weight of each community means that they are active in terms of content generation and impact, as well as interaction with hotel profiles on Twitter. Therefore, in Table 1 the main communities identified in terms of their weight and the themes of content shared by users are shown. We have also highlighted those communities that are directly linked to tourism-related content and that may involve improving decisions for CEOs and executives.

Table 1. Communities around digital tourism sorted by weight and activity

Community name	Weight	Community name	Weight
grasmere	305	**Pixabay***	**293**
derry	304	**Freestock***	**293**
government	303	mentoring	292
western	303	tradeup	292
Camping*	**302**	**Fitness***	**291**
Mountains*	**300**	weight	291
atomium	299	eventvenues	290
miniature	299	festivals	290
Accessibility*	**297**	**Summervacation***	**290**
government	303	**Summerfestival***	**290**
smokedmeat	296	merrimack	289
sandwich	296		
seniors	294		

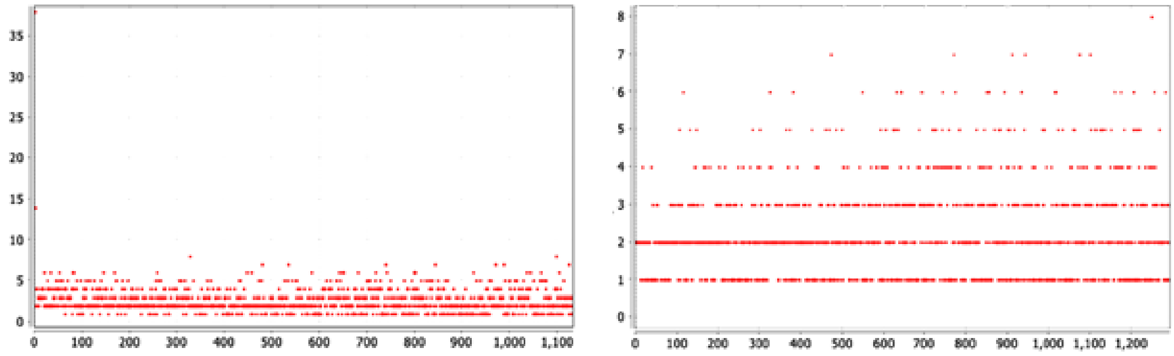
Source: Own Elaboration

Of the 24 communities identified, we highlight Camping, Mountains, Accessibility, Pizabay, Freestock, Fitness, Summervacation, and Summerfestival as those that are directly linked to the tourism sector to be defined, according to the UGC sample analyzed, some of the behaviors and preferences of tourists when sharing content on social networks.

The experiment for the visualization of data has been carried out with two different views and resolutions. On one hand, Figure 4 (a) shows the result for the modularity of 0.126 with a resolution of -0.020 that has resulted in a total number of communities of 1133. In Figure 4 (b) a modularity can be observed of 0.125 with a resolution of -0.022 and a total of 1291 user communities detected.

Although the literature indicates that these tests must be carried out to obtain a global vision of the subject analyzed in social networks, the authors have highlighted those communities that are most linked to tourism in Figure 4 according to their weight and modularity. In Figure 4, therefore, the diversity of themes and the number of communities analyzed with different resolutions can be observed, which allows us to observe that there are very specific UGC communities that are included in tourism in social networks.

Figure 4. Modularity report for digital tourism community's detection



(a) (b)

Source: Own Elaboration

5. CONCLUSION

After the development of the proposed research we can conclude that in terms of UGC on Twitter, there have been topics identified related to Camping, Mountains, Accessibility, Pixabay, Freestock, Fitness, Summervacation, and Summerfestival; these are most directly linked to user communities. Likewise, the same topics were also found to be related to other user communities which are not within the tourism sector (e.g., derry, government, western or Grasmere). We must consider that if the objective of companies in the tourism sector is to develop strategies that increase social network interactions, they should base their communications using the tag #instravel and tourism activities on the social network Twitter according to our exploratory experiments.

This fact allows the identification of strategies related to social media in the tourism sector, considering the instantaneity and publication of contents in real time. Results have been obtained to measure the number of UGC communities that exist around digital tourism on Twitter, taking into account their activity and weight.

These results show that tourists who have a travel destination in these or similar destinations are active on social networks, which can make it clear that companies take these communities as a reference to increase the impact of their social media strategies. Our analysis results verify that users usually interact on Twitter sharing information about campsites, photographs and, comments about mountains and accessibility. This fact links to sustainability in tourism, image banks in which content is shared, as well like sports and seasonal holiday seasons.

With respect to RQ1, it should be indicated that different algorithms have been used to visualize the communications and topics that characterize the UGC of tourism on the social network Twitter, providing insights related to social media strategies that can be used by CEOS and Executives for their tourism companies. Regarding RQ2, the most important topics have been offered as a result of the analysis when structuring and planning a content strategy for social media on Twitter, it does not lose value.

The implications for practitioners or managerial staff are of great interest since this study will allow them to correctly and more effectively plan social media content, make advertising investments on the Internet, and refine the choice of keywords as well as the topic that travelers choose for their social media publications.

Also, from the point of view of the methodology, our results help to understand how to interpret UGC data on Twitter and the main topics identified can be used by researchers as variables and constructs for their quantitative models, to test whether or not there exist positive meanings among them. Also, academics can use this research to better understand the tourism sector on social media and, in particular on Twitter, according to the type of content shared by users.

The limitations of this work lie in the exploratory nature of the research conducted. Our study is limited to comments on Twitter that link to hotels included in the Top 25 Hotels in the World as designated by TripAdvisor Travelers' Choice Awards. In addition, the results of the study are not significant enough to transfer to the whole digital tourism ecosystem. However, they continue the line of work that increasingly shows the importance of studying tweets and comments as a basis for detecting communities according to the type of content they share.

REFERENCES

Bennett, D., Yábar, D. P. B., & Saura, J. R. (2017). University Incubators May Be Socially Valuable, but How Effective Are They? A Case Study on Business Incubators at Universities. In Peris-Ortiz, M., Gómez, J., Merigó-Lindahl, J., Rueda-Armengot, C. (Eds.) *Entrepreneurial Universities. Innovation, Technology, and Knowledge Management*. Cham, Switzerland: Springer. doi: 10.1007/978-3-319-47949-1_11

Bizirgianni, I. & Dionysopoulou, P. (2013). The influence of tourist trends of youth tourism through social media (SM) & information and communication technologies (ICTs). *Procedia-Social and Behavioral Sciences*, 73, 652-660.

Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6), 1857-1864.

Confente, I., Siciliano, G. G., Gaudenzi, B., & Eickhoff, M. (2019). Effects of data breaches from user-generated content: A corporate reputation analysis. *European Management Journal.*

Corbin, J., Strauss, A., & Strauss, A. L. (2015). *Basics of qualitative research*. Sage. ISBN: 978-0803932517

Few, S. & Edge, P. (2007). *Data visualization: past, present, and future.* IBM Cognos Innovation Center.

Hays, S., Page, S. J., & Buhalis, D. (2013). Social media as a destination marketing tool: its use by national tourism organisations. *Current Issues in Tourism*, 16(3), 211-239.

Herráez, B., Bustamante, D., & Saura, J. R. (2017). Information classification on social networks. Content analysis of e-commerce companies on Twitter. *Revista Espacios*, 38(52), 16.

Jia, S. (2018). Leisure Motivation and Satisfaction: A Text Mining of Yoga Centres, Yoga Consumers, and Their Interactions. *Sustainability*, 10, 4458.

Jimenez-Marquez, J. L., Gonzalez-Carrasco, I., Lopez-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). Towards a big data framework for analyzing social media content. *International Journal of Information Management*, 44, 1-12.

Jovicic, D. Z. (2019). From the traditional understanding of tourism destination to the smart tourism destination. *Current Issues in Tourism*, 22(3), 276-282.

Kim, K. J. (2016). Interacting Socially with the Internet of Things (IoT): Effects of Source Attribution and Specialization in Human-IoT Interaction. *Journal of Computer-Mediated Communication*, 21(6), 420-435. doi:10.1111/jcc4.12177Murray 2015

Kucher, K. (2019). *Sentiment and Stance Visualization of Textual Data for Social Media* (Doctoral dissertation), Linnaeus University Press.

Leung, D., Law, R., Van Hoof, H., & Buhalis, D. (2013). Social media in tourism and hospitality: A literature review. *Journal of Travel & Tourism Marketing*, 30(1-2), 3-22.

Liu, H., Wu, L., & Li, X. (2019). Social media envy: how experience sharing on social networking sites drives millennials' aspirational tourism consumption. *Journal of Travel Research*, 58(3), 355-369.

Liu, Z., Jiang, B., & Heer, J. (2013). imMens: Real-time Visual Querying of Big Data. In *Computer Graphics Forum,* 32(3), 421-430. Oxford, UK: Blackwell Publishing Ltd.

Maimon, O. & Rokach, L. (Eds.) (2005). *Data mining and knowledge discovery handbook*. Springer.

Mkono, M. & Tribe, J. (2017). Beyond reviewing: Uncovering the multiple roles of tourism social media users. *Journal of Travel Research*, 56(3), 287-298.

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

Palos-Sanchez, P., Martin-Velicia, F., & Saura, J. R. (2018). Complexity in the Acceptance of Sustainable Search Engines on the Internet: An Analysis of Unobserved Heterogeneity with FIMIX-PLS. *Complexity*, 2018, 1-19. doi:10.1155/2018/6561417

Palos-Sanchez, P., Saura, J. R., & Martin-Velicia, F. (2019). A study of the effects of Programmatic Advertising on users' Concerns about Privacy overtime. *Journal of Business Research*, 96(2019), 61-72. doi: 10.1016/j.jbusres.2018.10.059

Reddy, C. S., Sangam, R. S., & Rao, B. S. (2019). A Survey on Business Intelligence Tools for Marketing, Financial, and Transportation Services. In *Smart Intelligent Computing and Applications* (pp. 495-504). Singapore: Springer.

Reyes-Menendez, A., Saura, J. R., & Palos-Sánchez, P. (2018). Crowdfunding y financiación 2.0. Un estudio exploratorio sobre el turismo cultural. *International Journal of Information Systems and Tourism (IJIST)*, 3(1), 23-34.

Reyes-Menendez, A., Saura, J. R., & Alvarez-Alonso, C. (2018). Understanding #WorldEnvironmentDay User Opinions in Twitter: A Topic-Based Sentiment Analysis Approach. *International Journal of Environmental Research and Public Health.* doi: doi.org/10.3390/ijerph15112537

Reyes-Menendez, A., Saura, J. R., Palos-Sanchez, P., & Alvarez, J. M. (2018). Understanding User Behavioral Intention to adopt a Search Engine that promotes Sustainable Water Management. *Symmetry-Basel*, 10(11), 584. doi: 10.3390/sym10110584

Reyes-Menendez, A., Saura, J. R., & Martinez-Navalon, J. G. (2019). The impact of e-WOM on Hotels Management Reputation: Exploring TripAdvisor Review Credibility with the ELM model. *IEEE Access*, 8(2). doi: 10.1109/ACCESS.2019.2919030

Saura, J. R., Palos-Sanchez, P. R., & Correia, M. B. (2019). Digital Marketing Strategies Based on the E-Business Model: Literature Review and Future Directions. In *Organizational Transformation and Managing Innovation in the Fourth Industrial Revolution* (pp. 86-103). IGI Global.

Saura, J. R., Palos-Sánchez, P., & Navalpotro, F. D. (2018). El problema de la Reputación Online y Motores de Búsqueda: Derecho al Olvido. *Cadernos de Dereito Actual*, 8, 221-229.

Saura, J. R., Reyes-Menendez, A., & Palos-Sanchez, P. (2018). Un Análisis de Sentimiento en Twitter con Machine Learning: Identificando el sentimiento sobre las ofertas de# BlackFriday. *Revista Espacios*, 39(16), 75.

Saura, J. R. & Bennett, D. (2019). A Three-Stage Methodological Process of Data Text Mining: A UGC Business Intelligence Analysis. *Symmetry-Basel*. doi: 10.13140/RG.2.2.11093.06880

Saura, J. R., Reyes-Menendez, A., & Palos-Sanchez, P. (2019). Are Black Friday Deals Worth It? Mining Twitter Users' Sentiment and Behavior Response. *Journal of Open Innovation: Technology, Market, and Complexity*, 5(3), 58. doi: 10.3390/joitmc5030058

Saura, J. R., Rodriguez Herráez, B., & Reyes-Menendez, A. (2019). Comparing a traditional approach for financial Brand Communication Analysis with a Big Data Analytics technique. *IEEE Access*, 7(1). doi: 10.1109/ACCESS.2019.2905301

Shaw, M. J., Subramaniam, C., Tan, G. W., & Welge, M. E. (2001). Knowledge management and data mining for marketing. *Decision Support Systems*, 31(1), 127-137.

Vassakis, K., Petrakis, E., Kopanakis, I., Makridis, J., & Mastorakis, G. (2019). Location-Based Social Network Data for Tourism Destinations. In *Big Data and Innovation in Tourism, Travel, and Hospitality* (pp. 105-114). Singapore: Springer.

Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51-65.

Zarco, C., Santos, E., & Cordón, O. (2019). Advanced visualization of Twitter data for its analysis as a communication channel in traditional companies. *Progress in Artificial Intelligence*, 1-17.