Review

A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality

Babajide Abubakr Muritala * [ID]**, Maria-Victoria Sánchez-Rebull** [ID] **and Ana-Beatriz Hernández-Lara** [ID]

Departament de Gestió d'Empreses, Facultat d'Economia i Empresa, Universitat Rovira i Virgili, 43204 Reus, Spain; mariavictoria.sanchez@urv.cat (M.-V.S.-R.); anabeatriz.hernandez@urv.cat (A.-B.H.-L.)

* Correspondence: babajide.muritala@urv.cat

check for updates

Abstract: This paper reviews the literature on online reviews in tourism and hospitality, and presents the current state of research in the area. A bibliometric approach was used to analyze 632 journal articles on online reviews in tourism and hospitality from 2005 to 2019 from the Scopus Database. This study identifies the most prolific journals, foundational works, and major research themes in the research area. In addition, we analyzed some dimensions of their network structure and the thematic evolution of the research area. The bibliometric method is quantitative and objective, and we carry out an analysis of the area based on citations and keywords. Researchers and business managers can gain useful insights on the current state of the art in this area. There have been only a few literature reviews tracking the growth in this research area, and even fewer using bibliometric methods or science maps. Therefore, this work provides an updated review of this fast-growing area with a bibliometric approach to highlight the recent developments with the aid of science maps, and shows the thematic network structure and evolution with an innovative visualization.

---

1. Introduction

Online reviews, also referred to as electronic word of mouth (eWOM) or user-generated content (UGC) [1,2], are all similar concepts with minor differences. eWOM is all electronic communications between producers and consumers and between consumers themselves: emails, websites, consumer review sites, blogs, virtual communities, chat rooms, newsgroups, and instant messaging [3]. UGC is data generated online by consumers, e.g., text (consumer reviews and blogs) and picture data [4].

Tourism and hospitality relies very much on word of mouth among consumers [5]. Before the Internet, this word of mouth was usually obtained from friends and acquaintances, but has now moved to the Internet, with consumers overwhelmingly consulting the opinions of fellow consumers online [6,7]. Their decision-making and travel experience is shaped by the reviews they read, and this has been a significant development in the industry [8].

Consumers use online reviews at every stage of the travel process—pre-trip, during trip, and post-trip—to engage in information search, share experiences, and give feedback [5]. Companies use it for customer engagement, to build online presence, get consumers' opinions, influence booking intentions, and earn revenue [9–11]. With its usefulness to both producers and consumers, it is no surprise that online reviews are a highly popular topic in tourism and hospitality for both researchers and managers [12].

As a result, there have been many studies on online reviews [13], but only a few keeping track of the literature. These works presented the state of the research at their time of publication, providing a

now-obsolete image of the field because of the astonishing speed at which the literature is growing. Therefore, recent studies on the application of online reviews in tourism and hospitality have not been integrated with earlier studies to understand the present state of research and current research trends. Many of these previous reviews have also been systematic [1,10] and none of them have used bibliometric methods or science maps [14] to visually show the development, conceptual structure, and thematic evolution of online reviews in tourism and hospitality.

Hence, the aim of this paper is to integrate and provide an organized summary of the existing research by identifying the foundational and seminal works and the old and new areas of the prior scholarship in order to build knowledge, gain an understanding, and show the future direction in this research area. We achieve this by means of bibliometric analytical techniques that enable us to deconstruct the main anchors and evolution of the research area. We also contribute an innovative way to show the thematic evolution of the key research themes. We draw four research questions based on this research objective to achieve this goal:

RQ1. What are the general trends in this research area?
RQ2. What is the foundational literature?
RQ3. What are the major research themes?
RQ4. How have the research interests evolved?

2. Literature Review

Even though research on online reviews in tourism and hospitality is fast-growing [14], it has a relatively brief history [15]. This is because notable online opinion platforms for customers rose to prominence in the early 2000s [12]. A few attempts have been made to review the literature and Table 1 presents 16 previous studies that we identified that reviewed the literature on online reviews in tourism and hospitality directly or indirectly.

Table 1 shows that these studies were published in reputable journals, with the International Journal of Contemporary Hospitality Management, Journal of Travel and Tourism Marketing, and International Journal of Hospitality Management having five, three, and two publications, respectively. The remaining journals—Tourism Management, Current Issues in Tourism, Journal of Hospitality Marketing and Management, Tourism Management Perspectives, International Journal of Hospitality and Tourism Administration, and Hospitality and Society—had one publication each.

An analysis of these previous reviews reveals that there are differences in their depth of focus. While some focused directly on online reviews in tourism and hospitality, many considered a larger domain such as social media or big data in tourism and hospitality, with online reviews treated as a subset of this larger domain.

Law et al. [16] had the biggest scope and explored information and communication technology (ICT) in tourism and hospitality, with online reviews included under ICT. This is followed in scope by Mariani et al. [17], who investigated business intelligence and big data in tourism and hospitality. Some studies reviewed research focused on types of big data in tourism and hospitality, including UGC such as online reviews, device data such as GPS (Global Positioning System) and Wi-Fi data, and transaction data such as web search data [4,14]. Other studies examined the role of social media in hospitality and tourism by reviewing the influence of online consumer review platforms, social networking sites, Internet forums, and other online communities [10,15,18–20]. Two studies examined eWOM relating only to hotels [1,21], whereas the remaining studies reviewed online reviews/UGC/eWOM in tourism and hospitality [2,12,13,22].

Table 1. Previous literature reviews.

Authors	Scope	Journal	Databases	Method	Sample Size	Years Covered
Leung et al. (2013)	SM [1] in T&H [2]	JTTM	SD [3], EBSCOHost, and GS [4]	Systematic	44	2007–2011
Serra Cantallops and Salvi (2014)	eWOM in Hotels	IJHM	6 selected journals	Systematic	28	2007–2011
Zeng and Gerritsen (2014)	SM in T&H	TMP	WoS [5], EBSCOHost, and GS	Bibliometric	279	2007–2013
Law et al. (2014)	ICT in T&H	IJCHM	SD, EBSCOHost, Emerald, and Sage	Content analysis	107	2009–2013
Schuckert et al. (2015)	Online reviews in T&H	JTTM	SD, EBSCOHost, and GS	Content analysis	50	2004–2013
Lu and Stepchenkova (2015)	UGC in T&H	JHMM	GS, WoS, top 5 tourism, and top 5 hospitality journals	Systematic	122	2001–2013
Chen and Law (2016)	eWOM in T&H	IJHTA	SD, EBSCOHost	Systematic	43	2008–2013
Kwok et al., (2017)	Online reviews in T&H	IJCHM	Top 7 T&H journals	Systematic	67	2000–2015
Sotiriadis (2017)	SM in T&H	IJCHM	SD, GS	Systematic	146	2009–2016
Bore et al. (2017)	eWOM in Hotels	HS	SD, EBSCOHost, Emerald, and SpringerLink	Systematic	45	2000–2015
Leung et al. (2017)	SM in T&H	IJHM	Top 8 business and top 8 T&H journals in WoS	Bibliometric	406	2007–2016
Li et al. (2018)	BD [6] in Tourism	TM	WoS, SD, Sage, Emerald, SpringerLink, Wiley Online Library, and GS	Systematic	165	2007–2017
Mariani et al. (2018)	BD in T&H	IJCHM	Scopus and WoS	Systematic	173	2000–2016
Lu et al. (2018)	SM in T&H	JTTM	SD, EBSCOHost, GS, and top 7 T&H journals	Systematic	105	2004–2014
Centobelli and Ndou (2019)	BD in Tourism	CIT	Scopus	Bibliometric	109	1990–2017
Nusair et al. (2019)	SM in T&H	IJCHM	SD and EBSCOHost	Bibliometric	439	2002–2016

[1] Social media; [2] Tourism and hospitality; [3] Science Direct; [4] Google Scholar; [5] Web of Science; [6] Big Data.

The previous works can be classified into three groups regarding the academic databases from which they selected their samples. The first group, which comprises most of the studies (11), used a keyword search of well-known databases such as Scopus and Google Scholar [4,10,13–18,21–23]. The second group concentrated on selected top journals such as *Tourism Management*, *Journal of Travel Research*, etc. [1,12,19]. The third group used both selected top journals and databases to get their samples [2,20].

Most of the studies considered only journal articles, except those of Li et al. [4], which used articles and conference papers; Centobelli and Ndou [14], which used articles, conference papers, and book chapters; and Zeng and Gerritsen [15], which used a lot of grey literature such as research degree theses, electronic articles, books, and reports in addition to articles and conference papers.

The literature reviews in this field started in 2013, and the year of publication explained the number of citations received by the documents, with the older publications having more citations than the newer ones, apart from some exceptions. Five of the 16 studies reviewed literature from 2007 up until their time of publication [1,4,10,15,19]. Three were from 2000 [12,17,21], two each started in 2004 [13,20] and 2009 [16,18], and one each in 2001 [2], 2002 [23] and 2008 [22]. Even though Centobelli and Ndou [14] searched for literature from 1990 to 2017, they reviewed documents published from 2011 onwards. In summary, the earliest time of publication of the reviewed literature was from the year 2000, with most of the studies considering literature from 2007 onwards. The last year considered in the most recent reviews was 2017 [4,14].

Regarding method, the previous reviews showed a preference for systematic reviews, with 10 of them adopting this approach [1,2,4,10,12,17,18,20–22]. Two applied content analysis [13,16] and four used a bibliometric approach [14,15,19,23]. The bibliometric reviews examined more documents, with three bibliometric reviews having the biggest sample sizes [15,19,23].

We employ the bibliometric analytical method for this study since it provides the tools to answer our research questions on the foundations and themes of this research area. The bibliometric method provides an advantage of objectivity and quantifiability, and helps to avoid subjective biases. It also helps to provide validation for findings that scholars had intuitively inferred in earlier studies and is more informative [24,25], which justifies the suitability of this methodological approach. Bibliometrics is the use of mathematical and statistical methods to quantify and analyze the bibliographic information of publications [26]. This bibliographic information enables researchers to make linkages between authors or papers [27]. Thomson Reuters Web of Science (WoS) and Elsevier's Scopus are the traditional databases for carrying out bibliometric studies since they are reliable sources of citation data [24].

Many bibliometric studies often utilize science maps [28]. The ability to visualize bibliometric networks with science maps was an important technical development in bibliometrics [24]. Science maps, also known as bibliometric maps or knowledge maps, help to reveal the conceptual, intellectual, or social structure of a field [28]. It is a spatial representation of the interrelationship between research elements such as authors or citations that facilitates the understanding of the structure and developments in a field [29]. There are different software tools for constructing science maps and Cobo et al. [28] made a comparative study of these tools. There are also different approaches to extracting a bibliometric network depending on the selected unit of analysis; e.g., authors, documents, journals, cited references, or keywords [28].

Science maps were used in just two of the four previous bibliometric reviews. Centobelli and Ndou [14] used a citation network, while Leung et al. [19] used a combination of co-citation and co-word analyses to make up for the weaknesses in a single method, and to reveal the theoretical foundation and thematic evolution of their area of study. This study follows in the steps of Leung et al. [19] with this combination for the same reason.

Co-citation analysis is a study of cited documents to get the frequency of citation of two earlier documents together. Hence, a co-citation link is a link between two documents cited together by a later document. When a set of documents is frequently co-cited together in a certain area, this may show

that the co-cited documents contain important concepts that peers have recognized. Thus, a co-citation network analysis can identify the core literature in a particular area of study [30].

Co-word or co-occurrence analysis measures the frequency of co-occurrence of two keywords in the same literature [31]. It relies on the assumption that the co-occurrence of keywords shows a non-trivial relationship [32]. When sets of keywords are frequently used together by different authors in a certain area, this may show that these keywords have a significant relationship within the research area [31]. This bibliometric method directly extracts the major themes of a research area and the linkages between these themes based on the co-occurrence of word pairs without relying on a priori definitions [33]. Hence, they help to reveal the conceptual structure and major research themes of the area [28,32].

The strength of a co-citation link is derived from the number of citing authors who cite two earlier works together. Therefore, co-citation is not a permanent relationship and the co-citation network pattern changes with time. The same applies to co-word analysis, which changes with a change in vocabulary co-occurrences in a field. Therefore, co-citation and co-word patterns in a field will change as the field evolves with a change in the interests and intellectual patterns [30,32].

From the review of the literature, it is obvious that the previous literature review studies have mainly been systematic reviews, and the few bibliometric studies covered a bigger scope, namely, social media [15,19,23] and big data [14]. Hence, their emphasis on online reviews was only peripheral. Similarly, the two previous bibliometric studies that used science maps focused on social media [19] and big data [14]. In summary, there have been no studies to the best of our knowledge providing the foundational literature and major research themes in the maturing research area of online reviews in tourism and hospitality. In addition, our study also covers the recent publications in a fast-growing field, as most of the documents reviewed in this work were published in the last two years. Finally, the thematic evolution chart and the analysis of some dimensions of its network structure is a novel contribution of this work that allows a better understanding of the development and state of the art in the field.

3. Method

3.1. Data Collection

We obtained the publications considered for this study from the Scopus database through a systematic process as shown in Figure 1. Scopus was chosen because it is a recognized index with a wide coverage of peer-reviewed publications and provides reliable bibliographic data [24,34]. Several search queries with relevant keywords were used (see Supplementary File S1) to search in the title, abstract, and keywords of the publications in the database from 2005 to 2019. The year 2005 was chosen as the start date because most of the previous reviews reported the first publications in this area to be from 2007 or later. Two years were added to this to ensure that all relevant documents were found. We filtered the results to include only publications in English and journal articles, resulting in 1440 documents. Their titles and abstracts were reviewed, and we selected 632 relevant publications. We selected theoretical and empirical articles on the application of online reviews in tourism and hospitality, with an emphasis on studies that used online review data and excluded technical papers focused on highlighting the performance and accuracy of algorithmic models for prediction, classification, recommendation engines, or detection of fake reviews.
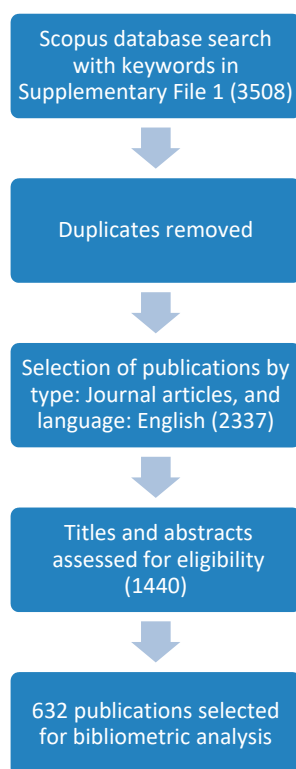
Figure 1. Systematic process to select literature.

3.2. Data Analysis

To answer the first research question posed by this study, the general literature trends are presented. Co-citation analysis was done to determine the foundational literatures of the study area, which are the most co-cited studies by the selected publications, and answers the second research question. A co-word analysis was also carried out because it enables the determination of the conceptual structure and research themes of the study area and helps to answer the third and fourth research questions. VOSviewer software (version 1.6.13) [35] and Ucinet 6.0 [36] were used to conduct the bibliometric analysis. Data cleaning, an important preprocessing step [28], was performed before data analysis by using a thesaurus file. The thesaurus file was used to combine variants of the same term or related concepts, such as "eWOM," "e-wom," "electronic word of mouth," etc., which were all merged into "eWOM." The thesaurus file was also used to ignore irrelevant words that did not provide any useful information, e.g., "article." The first step in the analysis was to produce a chart showing the growth in the number of published articles per year to show the growth in the literature. Next was a study of the most important journals, followed by the co-citation analysis to identify the core documents. Finally, the co-word analysis was carried out to identify the major research themes, with a study of the crucial dimensions of the keyword network structure and the evolution of the research interests.

4. Results

4.1. General Literature Trends

Figure 2 shows the growth in the number of publications within the study period. There was steady growth from 2006 to 2013 and a slight drop from 2013 to 2014, followed by rapid growth from 2014 onwards.
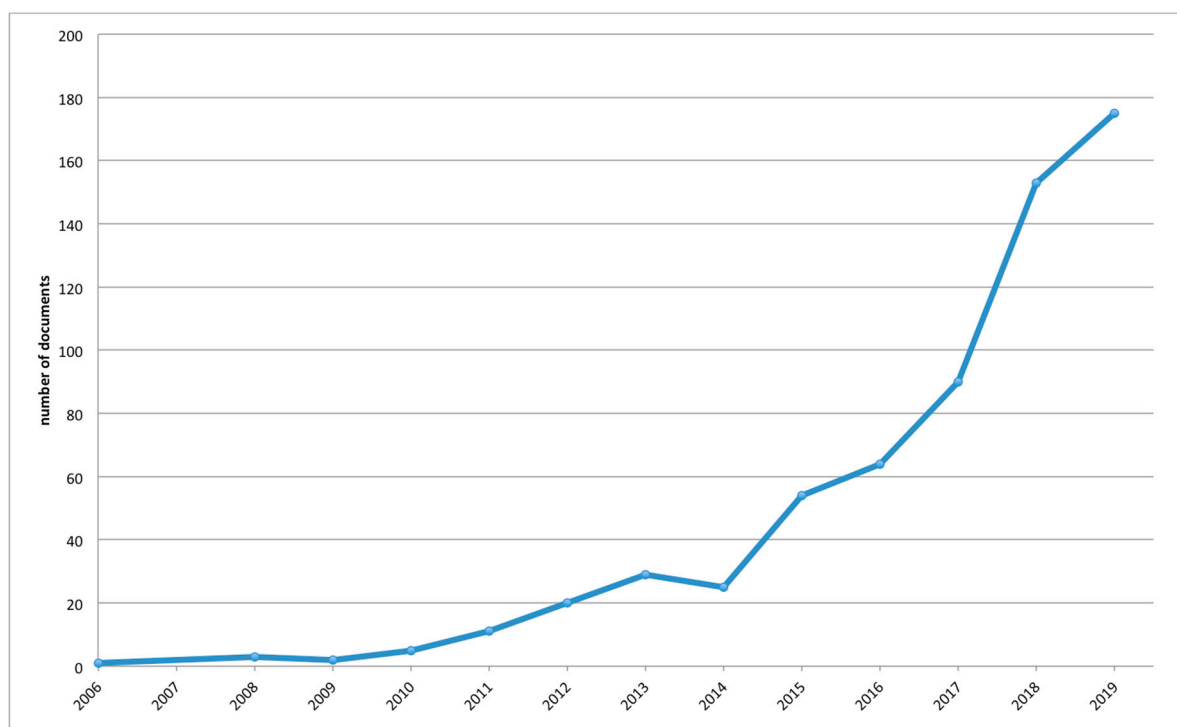
Figure 2. Growth of literature over time.

Table 2 shows the most prolific journals, using a threshold of six or more publications. The list includes many reputable tourism and hospitality journals, with most of them in the first quartile in the 2018 Scopus CiteScore under the Tourism, Leisure and Hospitality Management category. The 632 articles were from 202 journals.

Table 2. List of most prolific journals.

	Journal	Total Articles	Total Citations	Quartile
1	*Tourism Management*	56	3855	Q1
2	*International Journal of Hospitality Management*	53	2644	Q1
3	*International Journal of Contemporary Hospitality Management*	42	721	Q1
4	*Journal of Hospitality Marketing and Management*	25	675	Q1
5	*Current Issues in Tourism*	17	98	Q1
6	*Journal of Travel and Tourism Marketing*	17	885	Q1
7	*Sustainability (Switzerland)*	14	40	Q1 *
8	*Journal of Travel Research*	13	557	Q1
9	*Tourism Management Perspectives*	13	154	Q1
10	*Information Technology and Tourism*	12	70	Q1 *
11	*Journal of Hospitality and Tourism Technology*	12	165	Q1
12	*Cornell Hospitality Quarterly*	11	641	Q1
13	*Journal of Vacation Marketing*	9	150	Q1
14	*Journal of Hospitality and Tourism Research*	8	166	Q1
15	*Tourism Analysis*	8	15	Q3
16	*Asia Pacific Journal of Tourism Research*	7	115	Q1
17	*Journal of Business Research*	7	153	Q1 *

Table 2. *Cont.*

	Journal	Total Articles	Total Citations	Quartile
18	*Journal of China Tourism Research*	7	15	Q2
19	*Anatolia*	6	38	Q2 *
20	*Annals of Tourism Research*	6	131	Q1
21	*E Review of Tourism Research*	6	8	Q4
22	*International Journal of Culture Tourism and Hospitality Research*	6	26	Q2
23	*International Journal of Hospitality and Tourism Administration*	6	37	Q3
24	*Journal of Destination Marketing and Management*	6	165	Q1

Tourism Management and *International Journal of Hospitality Management* had the most publications; with both of them accounting for 17% of all publications. *Tourism Management* published articles with more impact, which have accumulated more citations than the others with about 26% of the total citations of all selected documents.

4.2. Foundational Literature

The co-citation network map was made to identify the foundational literature. To obtain this map, co-citation was selected as the type of analysis, cited references as the unit of analysis, and fractional counting as the counting method, as recommended [37]. Given the high number of cited references of over 30,000, an adequate minimum threshold of citations was mandatory for clarity. To obtain the top 20 most co-cited articles, as done in previous bibliometric studies [38], we set the threshold at 29 citations. The normalization method was association strength and the visualization weight was citations.

Figure 3 shows the co-citation network map. The proximity of two papers, the thickness of the lines connecting them, and their size shows the strength of their co-citation links [39]. These documents in the nodes are the 20 most influential or seminal papers in this area at this time. It is important to point out that even though these papers are not necessarily the most frequently cited papers in terms of bulk citations; they have been highly co-cited by the selected documents and hence represent the foundational literature.
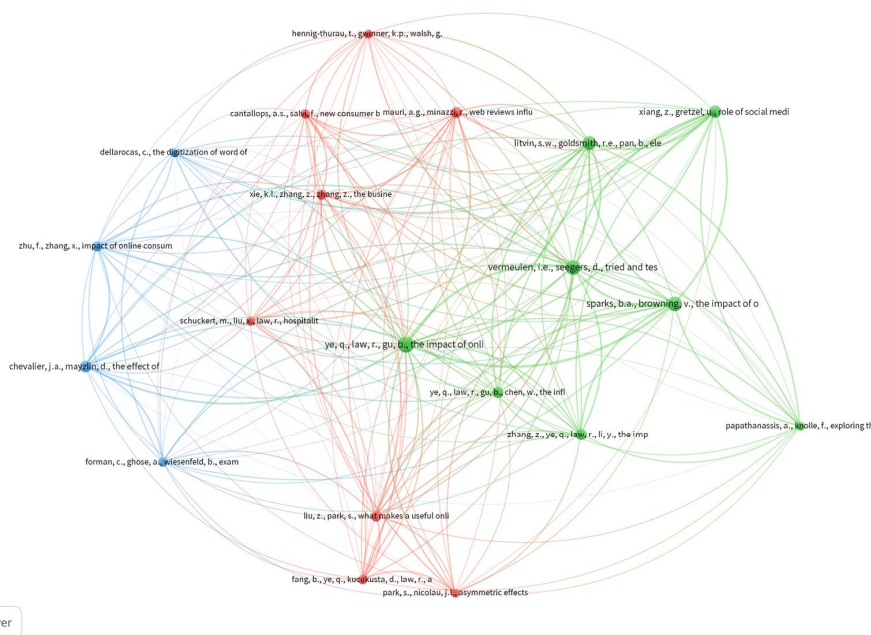

Figure 3. Co-citation network map.

Tourism Management and *International Journal of Hospitality Management* are the dominant sources of these influential papers, with seven and five publications, respectively, while other journals have one each of the remaining eight publications. These papers are in three clusters and there is a uniform pattern with the papers in each cluster. The blue cluster has four papers and are the earliest publications of the 20, published between 2003 and 2010. They deal with the effect of eWOM on sales and were written from outside the tourism and hospitality domain, specifically marketing, information systems, and management science [40–43].

The red cluster consists of eight papers. These are the relatively newer documents, and seven of them are the most recent documents of the 20, published between 2013 and 2016. They can be divided into three subgroups: The first two include literature reviews [1,13] and a paper on the motivations to write a review [44], whereas the remaining five papers are on the consequences of reviews, e.g., value and usefulness of online reviews [45–47], the impact of hotel reviews on consumer purchase intention and service expectations [48], and hotel performance [7].

The green cluster is also made up of eight papers and has larger nodes, which implies that they are the most co-cited papers on the map. They were published from 2008 to 2011 and can be situated in time between the early blue cluster and the recent red cluster. The papers in this mid-period were all written from a tourism and hospitality perspective, in contrast to the blue cluster. We can divide this cluster into two subgroups. The first is made up of influential conceptual papers on the role of social media in online travel information search [49], eWOM in hospitality and tourism [3], and an exploration of the adoption and processing of online reviews [50]. Although similar to some papers in the red cluster [7,48], the second subgroup comprises papers on the consequences of online reviews on different aspects of the sales process, such as consumer consideration [51], hotel booking and sales [9,52], and popularity of restaurants [53]. We find these similar papers in the red and green clusters relatively close on the visualization map.

4.3. Main Research Themes

The co-word network map was made to identify the major research themes. To obtain this map, co-occurrence was selected as the type of analysis, all keywords as the unit of analysis, and fractional counting as the counting method, and the thesaurus file was uploaded. The total number of keywords was 2028, which were too many to fit on a chart. Therefore, a threshold of three occurrences was set, which 219 keywords met. The normalization method was fractionalization and the visualization weight was occurrences.

Figure 4 shows the main keywords that have been used in online reviews literature (full list in Supplementary File S1). The proximity of two terms and the thickness of the lines connecting them show how frequently they co-occurred as keywords, and the size of a node is determined by the frequency of its occurrence as a keyword.

The most frequent keywords, such as "online reviews," "social media," "eWOM," "UGC," and "online ratings," had the biggest sizes. To avoid labels from overlapping, VOSviewer does not display labels close to a bigger label in the static image of the map, which prevents some keyword labels from showing. Hence, these oft-recurring keywords that cut across the entire field adding no extra information were removed from the visualization to simplify it and reveal the structure of the field spatially [33].

We calculated centrality measures to determine the precise structure of this network and identify the keywords that occupy critical positions within it. There are different centrality measures, all of which identify the most central elements in a network. Using Ucinet 6.0, degree, betweenness, and eigenvector centrality were calculated. These measures determine the prominence of a keyword in the network and existing conceptual overlap between them, however, some differences may also emerge from their use depending on the network configuration. Thus, considering all of them provides a more complete picture not just about the keywords that occupy highly central positions, but also

of the influence of these keywords on others [54] The top 10 central keywords for each measure are presented in Table 3, as done in previous studies [55].
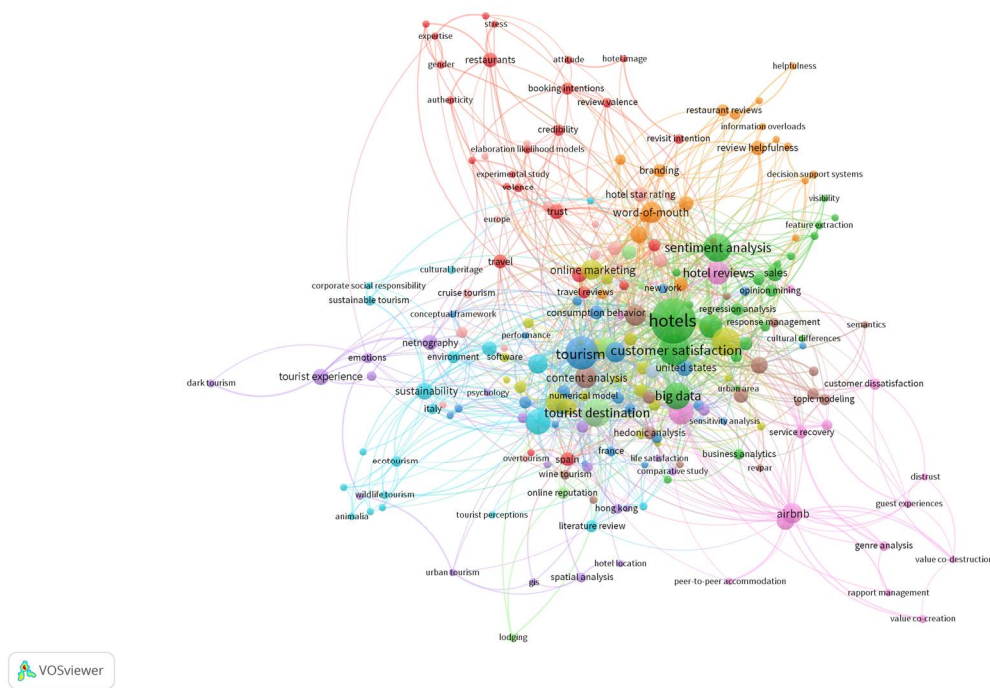

Figure 4. Co-word network map.

Table 3. Top 10 keywords with high centrality.

Keyword	Degree Centrality	Keyword	Betweenness Centrality	Keyword	Eigenvector Centrality
Hotels	96	Hotels	3100.50	Hotels	1.00
Tourism	53	Tourism	1453.66	Customer satisfaction	0.65
Customer satisfaction	45	Tourist destination	1240.57	Text analytics	0.61
Text analytics	43	Tourism management	1166.91	Tourism	0.60
Tourist destination	41	Customer satisfaction	1023.74	Sentiment analysis	0.51
Sentiment analysis	40	Hospitality	865.59	Big data	0.47
Tourism management	34	Text analytics	835.86	Hospitality	0.44
Big data	33	Sentiment analysis	767.95	Online ratings	0.42
Hospitality	33	Perception	644.48	Data mining	0.36
Data mining	28	Tourist behavior	584.82	Tourist destination	0.36

Hotels/hospitality (hotels, hotel reviews, hotel star rating, etc.) and tourism (tourism, tourist destination, tourism management, travel, travel reviews, wine tourism, cruise tourism, etc.) are the prominent industry domains and occupy a central position on the map, as expected according to all three centrality measures. Restaurants (restaurants, authenticity, booking intentions, etc.) was also an expected theme because of its relevance in the hospitality industry.

Customer satisfaction is also central, and highly connected on the map as a concept that has been investigated extensively using online reviews. Big data and text analytics techniques (sentiment analysis, topic modeling, etc.), and traditional research and statistical methodologies (regression

analysis, sensitivity analysis, content analysis, etc.) also occupy central positions as the processing tools for online reviews. It is worth mentioning that the high values of these keywords in terms of betweenness centrality underline not just their central position in the network, but also their influence on other keywords close to them.

Airbnb (Airbnb, sharing economy, peer-to-peer accommodation, guest experiences, value co-creation and co-destruction, etc.), sustainability (sustainability, sustainable tourism, ecotourism, overtourism, environment, wildlife tourism, corporate social responsibility, etc.), and tourist experience (emotions, dark tourism, heritage tourism, etc.) are also themes that can be identified on the map, showing the underlying structure of the research area.

4.4. Evolution of Research Interests

VOSviewer provides an overlay visualization of a network diagram, which shows the nodes on a network map with a color gradation based on the scores calculated by the software for either the average publication year (APY), average number of citations, or average normalized citations (ANC). To show the impact of the keywords over time in the selected literature in an easy-to-understand chart, the calculated figures from VOSviewer were taken and the keyword impact, represented by the ANC, was plotted against publication year (APY).

Figure 5 shows the wide range of keywords that have been used in the selected documents and the evolution of research interests in online reviews in tourism and hospitality in a timeline. ANC is the average number of normalized citations received by the documents that the keyword represents. The normalized citation count is the number of citations of the document divided by the average number of all citations of all documents published in the same year it was published, included in VOSviewer data [39]. ANC was used to compensate for the fact that citation counts favor older publications. APY is the average year of publication of the documents that the keyword represents. We placed the number of occurrences of the keywords in brackets in front of the keywords.
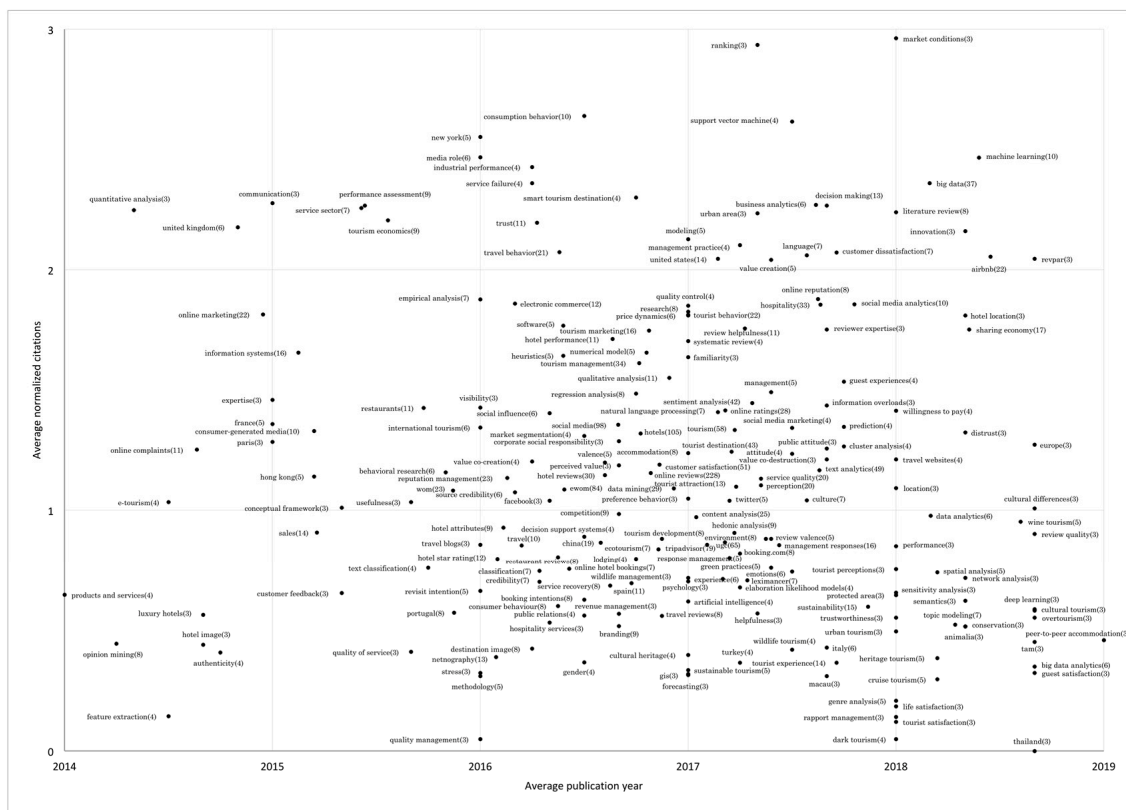

Figure 5. Evolution timeline of keywords.

This chart originally contained the 219 keywords from the co-word network map after choosing a threshold of three occurrences, with this threshold eliminating many keywords from earlier years. However, placing all 219 keywords on the chart, including those with very high ANC values and very low APY values, made it difficult to visualize most of the keywords. Therefore, five keywords with outlying values were excluded to get a clearer plot, with axes 0–3 for ANC and 2014–2019 for APY. We present these excluded keywords in Table 4. In addition, we present the full data of all 219 keywords in Supplementary File S2, i.e., ANC, APY, total number of occurrences, and the breakdown of occurrences by year.

Table 4. Outlier values removed from Figure 4.

Keyword	Occurrences	APY	ANC
Artificial neural network	4	2018.00	4.50
Tourism research	3	2016.67	3.65
Comparative study	4	2018.25	3.57
Experimental study	3	2013.67	1.80
Web 2.0	9	2013.78	1.81

Many of the keywords from the co-word analysis are visible in Figure 5, including those that could not be seen on the co-word network map to avoid overlapping. They include keywords about the industry sectors, consumer behavior, management processes to achieve management objectives, online review platforms, research methodologies, and country/city names. The closer to the right side of the plot the keyword appears, the more recent its occurrences in the literature are, and the closer to the top of the plot, the better cited (normalized citations) the articles are that used the keyword. Keywords found to the bottom right of the plot are emergent in the literature, while those at the top right are recent and well cited or hot topics [56].

The applications of artificial intelligence (AI) or machine learning techniques to the big data of online reviews are prominent themes in the hot topic area, e.g., "big data," "machine learning," "artificial neural network," "support vector machine," and "business analytics." Other keywords related to this theme, such as "deep learning," "text analytics," "natural language processing," "data analytics," "big data analytics," and "AI," are found at lower values of ANC, but still in the recent areas. This highlights their current relevance in the literature. The Airbnb and sharing economy theme is also in the hot topic area.

In the emergent area, online reviews are increasingly applied to investigate different niche areas of tourism, such as wine tourism, heritage tourism, cultural tourism, urban tourism, cruise tourism, wildlife tourism, and dark tourism. Sustainability themes such as sustainability, ecotourism, overtourism, environment, green practices, and sustainable tourism can also be found in this emergent area.

Figure 5 also shows a subtle evolution in terminology for some terms. For example, the average year of usage of "opinion mining" was 2014.3 (six of its eight appearances were between 2008 and 2016), and it seems to have now been replaced by terms such as "data mining" (2016.9), "text analytics" (2017.6), and "data analytics" (2018.2) in the tourism and hospitality literature. In the same vein, we can also observe an evolution in techniques. For example, feature extraction, a machine learning process used in text analysis to extract the unique features of a document [57], had an average year of 2014.5, and it seems to have now been eclipsed in the tourism and hospitality literature by topic modeling (2018.3), a method for carrying out feature extraction [57], which suggests that researchers now frequently use this technique for feature extraction. Another way the use of keywords has changed is the gradual change with time from the use of broader keywords to more specific sub-techniques in the literature as the research in the subfield deepens, for example, from a broad keyword like "AI" (2017) to "machine learning" (2018.4), a subset of AI, and then to "artificial neural networks" (2018) and "deep learning" (2018.67), which are both subsets of machine learning.

5. Discussion and Conclusions

This study used a bibliometric approach to review the research on online reviews in tourism and hospitality by analyzing 632 relevant documents selected from the Scopus database published between 2005 and 2019. The study applied co-citation and co-word analyses, including an innovative visualization of the co-word network to explore the theoretical foundations, network structure, and thematic evolution in this research area. It has successfully carried out an update, reviewed more studies, and revealed the theoretical foundations of the field with the use of science maps and shown the evolution of the major themes in the knowledge area. It has also provided an important quantitative complement with bibliometric methods to the previous literature reviews to enhance the understanding of this fast-growing knowledge area and has validated previous findings, such as the central research themes in online reviews.

Online review research is well established in the tourism and hospitality literature, with the top journals publishing literature in this research area. The earliest foundations of this research area was anchored to the adjacent body of knowledge on online reviews from other disciplines such as marketing and management, conducted on electronic commerce websites such as Amazon and eBay, before extensive research by tourism and hospitality scholars based on hotel and tourism reviews grew. Besides influential conceptual papers and literature reviews that integrate the knowledge base of the area, most of the foundational papers focused on the consequences, value, usefulness, and effect of online reviews on sales and financial performance, customer booking intention, and service expectations. These publications were important because they answered the fundamental question of whether online reviews were relevant to tourism and hospitality, because it is only when they are relevant that attention needs to be paid to them. Therefore, these papers played a crucial role as foundational works since they tackled the important theoretical concepts as well as the relevance, consequences, and challenges of eWOM. Hence, these documents were often co-cited by many publications in this research area to establish a theoretical base upon which to build their own research.

The co-word analysis confirms the research that has been carried out in the past in the traditional cornerstones of the tourism and hospitality industry: hotels, travel and tourism, and restaurants, as established in previous works [13,22], with customer satisfaction as a central concept and sentiment analysis as a predominant mode of analysis [4]. These themes not only occupy central positions but also maintain close relationships among themselves.

Big data and the application of machine learning are impacting virtually every industry [58]. The tourism and hospitality industry is not an exception, with the application of machine learning techniques such as artificial neural networks, support vector machines, and deep learning to natural language processing tasks, especially sentiment analysis. Many of these techniques were developed in the 1970s and 1980s, but big data and increased computer-processing capacity has resulted in their revival and wide application in different fields, including natural language processing [58,59]. We expect the application of these techniques to grow in the text processing of online reviews, especially in the deployment of novel sub-techniques as they are developed, as has happened in the past, i.e., in novel means of implementing neural networks and multilayer neural network (deep learning) models.

There is also an expansion of the application of online reviews beyond the traditional cornerstones of tourism and hospitality research into other themes, such as Airbnb, sustainability, and tourist experience. Airbnb's emergence with a unique accommodation business model has disrupted the tourism accommodation sector. The platform's celebration in March 2019 of reaching a milestone of 500 million guests, from about 10 million in June 2012, is evidence of its explosive growth [60,61]. Guest reviews are a feature of each Airbnb listing, and researchers have used these reviews to investigate guest experiences [62], host experiences [63], and other research perspectives, and is an active area of research in the application of online reviews to tourism and hospitality.

The theme of tourist experience, connected with emotions, heritage, and dark tourism, confirms the increasing demand for new forms of tourism and unique experiences, characterized by a departure from mass tourism [64] and an increase in demand for experiential tourism [65], which these niche

areas of tourism provide. The provision of memorable tourist experiences is directly related to a business's ability to generate revenue [66], and destination managers and even countries can leverage this. For example, Airbnb reported a five-fold increase in heritage travel globally since 2014 on its platform [67,68]. The application of online reviews to investigate these niche areas of tourism is likely to continue to grow, as there have been papers using online reviews to investigate other niche areas such as polar tourism [69], cycling tourism [70], plantation tours [71], etc. among the reviewed papers.

Tourism accounts for about 8% of global greenhouse gas emissions [72], and sustainability has become an important policy issue in tourism and hospitality [73]. Hence, "sustainability," "sustainable tourism," "ecotourism," and other related keywords are an important emergent area and reviews have been used to investigate guest reactions to hotel sustainability measures and the impact on guest experiences [74].

In conclusion, the research on online reviews and eWOM is reaching maturity with a rapidly increasing volume of new publications built on a solid base of seminal research and applied to highly varied areas, while increasingly sophisticated analytical methods are employed for analysis. This literature review has shown the foundational literature of this research area, recent growing areas of application in Airbnb and tourist experience, and in niche tourism areas such as wine tourism, dark tourism, etc. It has also revealed the latest analytical techniques being used, such as artificial neural networks and deep learning, and the hot and emergent topic areas, such as machine learning, big data analytics, and sustainability themes.

These findings are useful for scholars in this field to understand the current state of knowledge amidst the overwhelming growth in the literature and for general readers interested in learning about online reviews, and also provide practitioners and business owners a window to understanding a ubiquitous concept that they confront every day in the tourism and hospitality industry. The findings also provide a glimpse of future research growth areas toward the hot and emergent topics and the application of new machine learning sub-techniques using deep learning and neural networks for text processing. The thematic evolution chart, which highlights the hot and emergent topics, can be used in future bibliometric studies in any research area.

Limitations, Theoretical Implications, and Future Research

A limitation of this work is that authors sometime add arbitrary keywords to their papers that do not truly reflect the content of the paper, which could have affected the veracity of the co-word network map and thematic evolution timeline chart. The use of only the Scopus database is another limitation of this study that could have possibly resulted in leaving out a few publications indexed by WoS, but not in Scopus. It was not possible to combine both databases for this study because VOSviewer only allows the use of bibliographic data from either database, but not a combination of both. However, there is a significant overlap between both databases [34] and we recommend future work that can conduct a bibliometric analysis on a combination of both databases.

Our results have important implications for theory and future research. They show the conceptual structure and thematic evolution of this research area, and an individual or researcher new to this research area can quickly obtain a bird's-eye view of the seminal publications and how the research themes have evolved to date. However, research is a continuous process and as mentioned, the co-citation network map, co-word network map, and, by extension, the thematic evolution chart are not static but change with time. Future studies are thereby needed to track the changes that take place in the conceptual structure and major research themes of this research area with time. In addition, in light of the COVID-19 pandemic, studies employing online reviews and eWOM from various sources such as social media are suggested to understand the various effects of the pandemic.

References

1. Serra Cantallops, A.; Salvi, F. New consumer behavior: A review of research on eWOM and hotels. *Int. J. Hosp. Manag.* **2014**, *36*, 41–51. [CrossRef]

2. Lu, W.; Stepchenkova, S. User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software. *J. Hosp. Mark. Manag.* **2015**, *24*, 119–154. [CrossRef]

3. Litvin, S.W.; Goldsmith, R.E.; Pan, B. Electronic word-of-mouth in hospitality and tourism management. *Tour. Manag.* **2008**, *29*, 458–468. [CrossRef]

4. Li, J.; Xu, L.; Tang, L.; Wang, S.; Li, L. Big data in tourism research: A literature review. *Tour. Manag.* **2018**, *68*, 301–323. [CrossRef]

5. Fotis, J.; Buhalis, D.; Rossides, N. *Social Media Use and Impact during the Holiday Travel Planning Process BT—Information and Communication Technologies in Tourism 2012*; Fuchs, M., Ricci, F., Cantoni, L., Eds.; Springer Vienna: Vienna, Austria, 2012; pp. 13–24.

6. Wang, Y.; Yu, Q.; Fesenmaier, D.R. Defining the virtual tourist community: Implications for tourism marketing. *Tour. Manag.* **2002**, *23*, 407–417. [CrossRef]

7. Xie, K.L.; Zhang, Z.; Zhang, Z. The business value of online consumer reviews and management response to hotel performance. *Int. J. Hosp. Manag.* **2014**, *43*, 1–12. [CrossRef]

8. Sparks, B.A.; Perkins, H.E.; Buckley, R. Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tour. Manag.* **2013**, *39*, 1–9. [CrossRef]

9. Sparks, B.A.; Browning, V. The impact of online reviews on hotel booking intentions and perception of trust. *Tour. Manag.* **2011**, *32*, 1310–1323. [CrossRef]

10. Leung, D.; Law, R.; van Hoof, H.; Buhalis, D. Social Media in Tourism and Hospitality: A Literature Review. *J. Travel Tour. Mark.* **2013**, *30*, 3–22. [CrossRef]

11. He, W.; Tian, X.; Tao, R.; Zhang, W.; Yan, G.; Akula, V. Application of social media analytics: A case of analyzing online hotel reviews. *Online Inf. Rev.* **2017**, *41*, 921–935. [CrossRef]

12. Kwok, L.; Xie, K.L.; Richards, T. Thematic framework of online review research: A systematic analysis of contemporary literature on seven major hospitality and tourism journals. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 307–354. [CrossRef]

13. Schuckert, M.; Liu, X.; Law, R. Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *J. Travel Tour. Mark.* **2015**, *32*, 608–621. [CrossRef]

14. Centobelli, P.; Ndou, V. Managing customer knowledge through the use of big data analytics in tourism research. *Curr. Issues Tour.* **2019**, *22*, 1862–1882. [CrossRef]

15. Zeng, B.; Gerritsen, R. What do we know about social media in tourism? A review. *Tour. Manag. Perspect.* **2014**, *10*, 27–36. [CrossRef]

16. Law, R.; Buhalis, D.; Cobanoglu, C. Progress on information and communication technologies in hospitality and tourism. *Int. J. Contemp. Hosp. Manag.* **2014**, *26*, 727–750. [CrossRef]

17. Mariani, M.; Baggio, R.; Fuchs, M.; Höepken, W. Business intelligence and big data in hospitality and tourism: A systematic literature review. *Int. J. Contemp. Hosp. Manag.* **2018**, *30*, 3514–3554. [CrossRef]

18. Sotiriadis, M.D. Sharing tourism experiences in social media: A literature review and a set of suggested business strategies. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 179–225. [CrossRef]

19. Leung, X.Y.; Sun, J.; Bai, B. Bibliometrics of social media research: A co-citation and co-word analysis. *Int. J. Hosp. Manag.* **2017**, *66*, 35–45. [CrossRef]

20. Lu, Y.; Chen, Z.; Law, R. Mapping the progress of social media research in hospitality and tourism management from 2004 to 2014. *J. Travel Tour. Mark.* **2018**, *35*, 102–118. [CrossRef]

21. Bore, I.; Rutherford, C.; Glasgow, S.; Aheri, B.; Antony, J. A systematic literature review on eWOM in the hotel industry: Current trends and suggestions for future research. *Hosp. Soc.* **2017**, *7*, 63–85. [CrossRef]

22. Chen, Y.-F.; Law, R. A Review of Research on Electronic Word-of-Mouth in Hospitality and Tourism Management. *Int. J. Hosp. Tour. Adm.* **2016**, *17*, 347–372. [CrossRef]

23. Nusair, K.; Butt, I.; Nikhashemi, S.R. A bibliometric analysis of social media in hospitality and tourism research. *Int. J. Contemp. Hosp. Manag.* **2019**, *31*, 2691–2719. [CrossRef]

24. Mingers, J.; Leydesdorff, L. A review of theory and practice in scientometrics. *Eur. J. Oper. Res.* **2015**, *246*, 1–19. [CrossRef]

25. Nerur, S.P.; Rasheed, A.A.; Natarajan, V. The intellectual structure of the strategic management field: An author co-citation analysis. *Strateg. Manag. J.* **2008**, *29*, 319–336. [CrossRef]

26. Pritchard, A. Statistical bibliography or bibliometrics? *J. Doc.* **1969**, *25*, 348–349.

27. Garfield, E.; Malin, V.M.; Small, H. Citation Data as Science Indicators. *Towar. A Metr. Sci.* **1983**, *6*, 580.

28. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. Science mapping software tools: Review, analysis, and cooperative study among tools. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 1382–1402. [CrossRef]

29. Small, H. Visualizing science by citation mapping. *J. Am. Soc. Inf. Sci.* **1999**, *50*, 799–813. [CrossRef]

30. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [CrossRef]

31. Whittaker, J. Creativity and Conformity in Science: Titles, Keywords and Co-word Analysis. *Soc. Stud. Sci.* **1989**, *19*, 473–496. [CrossRef]

32. Callon, M.; Courtial, J.P.; Laville, F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics* **1991**, *22*, 155–205. [CrossRef]

33. He, Q. Knowledge Discovery Through Co-Word Analysis. *Libr. Trends* **1999**, *48*, 133–159.

34. Martín-Martín, A.; Orduna-Malea, E.; Thelwall, M.; Delgado López-Cózar, E. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *J. Informetr.* **2018**, *12*, 1160–1177. [CrossRef]

35. van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef]

36. Borgatti, S.P.; Everett, M.G.; Freeman, L.C. *UCINET 6 for Windows: Software for Social Network Analysis*; Analytic Technologies: Harvard, MA, USA, 2002.

37. van Eck, N.J.; Waltman, L. *Visualizing Bibliometric Networks*; Springer: Cham, Switzerland, 2014; ISBN 9783319103778.

38. Liu, Z.; Yin, Y.; Liu, W.; Dunford, M. Visualizing the intellectual structure and evolution of innovation systems research: A bibliometric analysis. *Scientometrics* **2015**, *103*, 135–158. [CrossRef]

39. van Eck, N.J.; Waltman, L. *VOSviewer Manual*; Univeristeit Leiden: Leiden, The Netherlands, 2013.

40. Chevalier, J.A.; Mayzlin, D. The effect of word of mouth on sales: Online book reviews. *J. Mark. Res.* **2006**, *43*, 345–354. [CrossRef]

41. Dellarocas, C. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Manag. Sci.* **2003**, *49*, 1407–1424. [CrossRef]

42. Forman, C.; Ghose, A.; Wiesenfeld, B. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inf. Syst. Res.* **2008**, *19*, 291–313. [CrossRef]

43. Zhu, F.; Zhang, X. Impact of online consumer reviews on Sales: The moderating role of product and consumer characteristics. *J. Mark.* **2010**, *74*, 133–148. [CrossRef]

44. Hennig-Thurau, T.; Gwinner, K.P.; Walsh, G.; Gremler, D.D. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *J. Interact. Mark.* **2004**, *18*, 38–52. [CrossRef]

45. Fang, B.; Ye, Q.; Kucukusta, D.; Law, R. Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tour. Manag.* **2016**, *52*, 498–506. [CrossRef]

46. Liu, Z.; Park, S. What makes a useful online review? Implication for travel product websites. *Tour. Manag.* **2015**, *47*, 140–151. [CrossRef]

47. Park, S.; Nicolau, J.L. Asymmetric effects of online consumer reviews. *Ann. Tour. Res.* **2015**, *50*, 67–83. [CrossRef]

48. Mauri, A.G.; Minazzi, R. Web reviews influence on expectations and purchasing intentions of hotel potential customers. *Int. J. Hosp. Manag.* **2013**, *34*, 99–107. [CrossRef]

49. Xiang, Z.; Gretzel, U. Role of social media in online travel information search. *Tour. Manag.* **2010**, *31*, 179–188. [CrossRef]

50. Papathanassis, A.; Knolle, F. Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tour. Manag.* **2011**, *32*, 215–224. [CrossRef]

51. Vermeulen, I.E.; Seegers, D. Tried and tested: The impact of online hotel reviews on consumer consideration. *Tour. Manag.* **2009**, *30*, 123–127. [CrossRef]

52. Ye, Q.; Law, R.; Gu, B.; Chen, W. The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Comput. Hum. Behav.* **2011**, *27*, 634–639. [CrossRef]

53. Zhang, Z.; Ye, Q.; Law, R.; Li, Y. The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *Int. J. Hosp. Manag.* **2010**, *29*, 694–700. [CrossRef]

54. Valente, T.; Coronges, K.; Lakon, C.; Costenbader, E. How Correlated Are Network Centrality Measures? *Connect. (Tor)* **2008**, *28*, 16–26.

55. Hu, C.-P.; Hu, J.-M.; Deng, S.-L.; Liu, Y. A co-word analysis of library and information science in China. *Scientometrics* **2013**, *97*, 369–382. [CrossRef]

56. Ávila-Robinson, A.; Wakabayashi, N. Changes in the structures and directions of destination management and marketing research: A bibliometric mapping study, 2005–2016. *J. Destin. Mark. Manag.* **2018**, *10*, 101–111. [CrossRef]

57. Liu, B. *Sentiment Analysis and Opinion Mining*; Morgan & Claypool Publishers: San Rafael, CA, USA, 2012; Volume 5.

58. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]

59. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]

60. Guttentag, D. Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Curr. Issues Tour.* **2015**, *18*, 1192–1217. [CrossRef]

61. Airbnb Airbnb Celebrates Half A Billion Guest Arrivals. Available online: https://news.airbnb.com/airbnb-celebrates-half-a-billion-guest-arrivals/ (accessed on 21 December 2019).

62. Cheng, M.; Jin, X. What do Airbnb users care about? An analysis of online review comments. *Int. J. Hosp. Manag.* **2019**, *76*, 58–70. [CrossRef]

63. Cheng, M.; Zhang, G. When Western hosts meet Eastern guests: Airbnb hosts' experience with Chinese outbound tourists. *Ann. Tour. Res.* **2019**, *75*, 288–303. [CrossRef]

64. Stamboulis, Y.; Skayannis, P. Innovation strategies and technology for experience-based tourism. *Tour. Manag.* **2003**, *24*, 35–43. [CrossRef]

65. Smith, W.L. Experiential tourism around the world and at home: Definitions and standards. *Int. J. Serv. Stand.* **2006**, *2*, 1–14. [CrossRef]

66. Kim, J.-H.; Ritchie, J.R.B.; McCormick, B. Development of a scale to measure memorable tourism experiences. *J. Travel Res.* **2012**, *51*, 12–25. [CrossRef]

67. Airbnb Heritage Travel on the Rise: Airbnb and 23andMe Team Up to Make it Even Easier. Available online: https://news.airbnb.com/heritage-travel-on-the-rise/ (accessed on 19 December 2019).

68. Yearofreturn.com Year Of Return—Celebrating 400 years of African Resilience. Available online: https://www.yearofreturn.com/ (accessed on 19 December 2019).

69. Heimtun, B. Emotions and affects at work on Northern Lights tours. *Hosp. Soc.* **2016**, *6*, 223–241. [CrossRef]

70. Chiu, W.; Leng, H.K. Let's go cycling: An analysis of tourists' experience on online user-generated content. *Int. J. Tour. Cities* **2017**, *3*, 30–42. [CrossRef]

71. Carter, P.L. Where are the enslaved?: Tripadvisor and the narrative landscapes of southern plantation museums. *J. Herit. Tour.* **2016**, *11*, 235–249. [CrossRef]

72. Lenzen, M.; Sun, Y.-Y.; Faturay, F.; Ting, Y.-P.; Geschke, A.; Malik, A. The carbon footprint of global tourism. *Nat. Clim. Chang.* **2018**, *8*, 522–528. [CrossRef]

73.   Saarinen, J. Traditions of sustainability in tourism studies. *Ann. Tour. Res.* **2006**, *33*, 1121–1140. [CrossRef]
74.   Gerdt, S.-O.; Wagner, E.; Schewe, G. The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tour. Manag.* **2019**, *74*, 155–172. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.