

Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Mariani, Marcello ORCID logoORCID: <https://orcid.org/0000-0002-7916-2576> and Borghi, Matteo ORCID logoORCID: <https://orcid.org/0000-0002-4150-1595> (2023) Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings. *Journal of Sustainable Tourism*, 31 (11). pp. 2592-2611. ISSN 1747-7646 doi: <https://doi.org/10.1080/09669582.2022.2033982> Available at <https://centaur.reading.ac.uk/103848/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1080/09669582.2022.2033982>

Publisher: Informa UK Limited

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings

Marcello Mariani & Matteo Borghi

To cite this article: Marcello Mariani & Matteo Borghi (2022): Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings, Journal of Sustainable Tourism, DOI: [10.1080/09669582.2022.2033982](https://doi.org/10.1080/09669582.2022.2033982)

To link to this article: <https://doi.org/10.1080/09669582.2022.2033982>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 17 Feb 2022.



Submit your article to this journal [↗](#)



Article views: 591



View related articles [↗](#)



View Crossmark data [↗](#)

Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings

Marcello Mariani*  and Matteo Borghi

Henley Business School, University of Reading, Henley-on-Thames, Greenlands, Oxfordshire, United Kingdom

ABSTRACT

By deploying big data analytical techniques to retrieve and analyze a large volume of more than 2.7 million reviews, this work sheds light on how environmental concerns expressed by tourists on digital platforms, in the guise of online reviews, influence their satisfaction with tourism and hospitality services. More specifically, we conduct a multi-platform study of Tripadvisor.com and Booking.com online reviews (ORs) pertaining to hotel services across eight leading tourism destination cities in America and Europe over the period 2017–2018. By adopting multivariate regression analyses, we show that OR ratings are positively influenced by both the *presence* and *depth* of environmental discourse on these platforms. Theoretical and managerial contributions, and implications for digital platforms, big data analytics (BDA), electronic word-of-mouth (eWOM) and environmental research within the tourism and hospitality domain are examined, with a view to capturing, empirically, the effect of environmental discourse *presence* and *depth* on customer satisfaction proxied through online ratings.

ARTICLE HISTORY

Received 23 February 2021
Accepted 20 January 2022

KEYWORDS

Big data analytics;
Booking.com; digital
platforms; eWOM;
environmental concerns;
online review platforms;
Tripadvisor

Introduction

Human consumption activities have been found to generate natural resources depletion and jeopardize the very existence of the planet (United Nations, 2013). Tourism is no exception. As a subset of human activities, tourism-related activities have been found to consume considerable natural resources. In 2010, it was estimated that tourism-related activities consumed (on average) 3,575MJ of energy per trip, 6,575 liters of water per tourist, 42 m² of land per bed, 1,800 g of food per tourist/day, and generate 250 kg CO₂ of emission per trip (Gössling & Peeters, 2015). The overall situation is getting worst over time, with growth factors over the period 2010–2050 estimated to be 2.64 for energy and CO₂ emissions, 1.92 for fresh water, 2.89 for land use and 2.08 for food use. Tourism activities – the tourism and hospitality infrastructure, and transit activities – have a detrimental impact on the environment (for example, in the form of CO₂ emission). An excess of tourism flow, known as over-tourism, is identified as one of the major issues associated with tourism activities and has been increasingly researched (Oklevik et al., 2019).

CONTACT Marcello Mariani  m.mariani@henley.ac.uk  Henley Business School, University of Reading, Henley-on-Thames, Greenlands, Oxfordshire RG9 3AU, United Kingdom.

*University of Bologna, Bologna, Italy.

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Digital technologies in the hospitality and tourism domains are bringing about both risk and opportunities for society. As these technologies empower consumers, they can engender further tourism over-consumption (Gössling, 2017), but they can also make business processes more environmentally friendly and help monitor consumption levels in real time (Gorgemans & Murillo-Luna, 2016). Digital technologies and user-generated content (UGC) have become “the single most important new determinant in tourism’s demand and supply structures” (Gössling, 2017, p.1025). Accordingly, and as recognized by recent studies (Mariani & Borghi, 2021a), digital technologies and digitalization have significantly impacted on the demand of tourism, changing the way consumers book and experience tourism products, and especially how consumers seek sustainable tourism products, services and experiences. However, so far, tourism and hospitality research has typically examined traditional media coverage of environment-related issues, largely disregarding the deployment of UGC-derived big data (BD) and analytics to generate insights on consumers’ opinions of environmental issues. This is a major knowledge gap, originally raised in a recent study (Mariani & Borghi, 2021a), that we intend to address with the present work. Accordingly, we argue that travel-related online reviews (ORs) and UGC – and the related analytics – can be leveraged to detect, monitor and analyze travelers’ and consumers’ environmental perceptions and concerns, therefore moving beyond their use to capture consumers’ satisfaction with hospitality and tourism services (Xiang et al., 2017). As such, ORs can proxy online consumers’ environmental concerns about tourism and hospitality services and products. More specifically, we build on the established and renowned marketing concept of electronic word-of-mouth (eWOM) (Hennig-Thurau et al., 2004) and use the conceptualization of online consumers’ environmental discourse defined as “eWOM in the form of ORs directly related to online consumers’ evaluations of environmental issues” (Mariani & Borghi, 2021a, p.833) and the distinction between the *presence* and *depth* of online consumers’ environmental discourse to address the following research question: “What is the effect of online consumers’ environmental discourse *presence* and *depth* on online review ratings?”

To address this research question and overcome the empirical limitations of previous studies using convenience samples of customers to make sense of the social and environmental impact on customer satisfaction (Brazyté et al., 2017; D’Acunto et al., 2020; Ettinger et al., 2018; Lee et al., 2016; Yu et al., 2017), we deploy advanced (big) data analysis techniques to retrieve and analyze the entire population of Tripadvisor and Booking.com ORs, covering hospitality services in eight leading cities renowned for tourism in America (Las Vegas, Miami, New York City and Orlando) and Europe (London, Paris, Rome and Barcelona) over a long time. Building on more than 2.7 million ORs pertaining to hotel services, delivered by an unparalleled number of hotel companies in environmental studies (5,572 and 5,044 hotels listed on Tripadvisor and Booking.com, respectively), this study uses multivariate regression analyses to examine the effect of online consumers’ environmental discourse *presence* and *depth* on online review ratings across the two platforms: a community-based review platform (i.e. Tripadvisor) vs. a transaction-based online travel agency (i.e. Booking.com).

As such, the study contributes to a nascent research stream at the intersection of digital platforms, eWOM, big data analytics (BDA) and sustainable tourism. Here we summarize three contributions and provide the full list of seven contributions in our discussion and conclusion. First, we contribute to sustainable tourism research by addressing how the *presence* and *depth* of online consumers’ environmental discourse emerge on digital platforms, thus complementing a nascent research line (Mariani & Borghi, 2021a) which has simply tracked longitudinally the *presence* and *depth* of online consumers’ environmental discourse. Our study differs from previous literature – and especially the work by Mariani and Borghi (2020) – in that it quantitatively measures how consumers’ environmental concerns influence online consumer satisfaction with tourism and hospitality services. Secondly, we find that tourists seek sustainable experiences (Oklevik et al., 2019) in online settings, and voice their environmental concerns on digital platforms through ORs. Thus, ORs are not only a rich data source to gain insights into online

consumers' environmental discourse, but they also allow researchers to understand – by means of BDA – how virtual customers (Nambisan & Baron, 2007; Nambisan & Nambisan, 2008) engaging with sustainability, can help generate important insights about prospective customers' behaviors. Third, by focusing on the online tourists' environmental concerns, this work enriches the body of eWOM literature in the sustainable tourism field that, so far, has mainly focused on broad corporate social responsibility (CSR) discourses (e.g. Ettinger et al., 2018) or green practices at the local or, at best, national level (e.g. Lee et al., 2016; Yu et al., 2017). More specifically, we enrich the research stream at the intersection between online consumers' perceptions of environmental issues through eWOM and their satisfaction with hospitality and tourism services (e.g. Brazyté et al., 2017; Lee et al., 2016; Yu et al., 2017), by finding that there is a positive relationship between environmental discourse presence/depth and online customer satisfaction.

Theoretical background and literature

Digital technologies, digital platforms and big data analytics in tourism

Digital technologies have been recognized as major drivers of the ongoing digital revolution (Rüßmann et al., 2015) and digital platforms the means through which novel business models are developed (UNCTAD., 2019) and value is created in the platform economy (Kenney & Zysman, 2016; Mariani and Nambisan, 2021; Mariani et al., 2021). Big data and analytics have been recognized as one of the technological drivers of the digital revolution and transformation of business. Beyond representing a technological paradigm per se, BD has been demarcated and conceptualized as a vast amount of data, whether structured or unstructured, which is produced at high speed because of technological advancements and the growth and diffusion of automation, the internet and connected devices (Mariani et al., 2018).

Over the last decade an increasing number of scientists, scholars and practitioners are relying on BD and analytics to uncover patterns in data which can translate into competitive business intelligence (Davenport, 2014) as well as knowledge (Erevelles et al., 2016). Also, scholars and practitioners working in the travel and tourism domains are deploying BD and analytics (Li et al., 2018; Mariani et al., 2018). BD and analytics are generated from various sources encompassing transactions, devices and UGC. Transaction data include online booking data, consumer cards data and web search data; more specifically web search data was deployed to forecast tourism demand (Pan et al., 2011; Peng et al., 2017). Device data encompass GPS, mobile roaming, Bluetooth, radio-frequency identification (RFID); more specifically, travelers' location data have generated useful insights about travel patterns and purpose (Gong et al., 2016), transportation modes (Kasahara et al., 2015) and points of tourism interest. Yet, UGC is certainly the most common source of data – for instance, ORs have been widely used to capture online customer satisfaction and consumers' evaluations of travel experiences, hospitality services (Guo et al., 2017; Xiang et al., 2017) and tourism destinations (Mariani et al., 2018). Geo-tagged online content, such as photos and social media posts, have been able to reveal the geographic distribution of travelers and residents, as well as tourism attractions with the most positive destination image (e.g. Marine-Roig & Clavé, 2015; Paldino et al., 2015; Salvatella, 2014).

Big data analytics is a holistic process: patterns are extracted from BD and generate actionable insights and create business value (Fosso Wamba et al., 2020; Mariani & Fosso Wamba, 2020; Mariani & Nambisan, 2021). As part of this process, several techniques are being increasingly used to make sense of users' behaviors in the tourism and hospitality sector. For instance, tourism and hospitality scholars have deployed techniques such as sentiment analysis and analytics (e.g. Alaei et al., 2019; Hao et al., 2020; Mehraliyev et al., 2020; Serrano et al., 2021), textual meta-data processing (e.g. Miah et al., 2017), machine learning techniques (e.g. Ahani et al., 2019; Sun et al., 2019; Xiang et al., 2017), deep learning models (e.g. Chang et al., 2020; Ma et al., 2018; Zhang et al., 2020), and topic modeling techniques (Mirzaalian & Halpenny, 2019; Moro et al.,

2019; Sutherland et al., 2020) to make sense of behaviors within tourism and hospitality. These techniques have been adopted in a wide range of disciplines, including information management, information systems and marketing (France & Ghose, 2019; Gandomi & Haider, 2015; Wedel & Kannan, 2016).

User-generated content, in the form of social media content, posts and online reviews, is increasingly relevant not only to understand more about consumers' opinions and preferences about travel, tourism and hospitality products, but also to generate business intelligence to improve products, services and attractions (Fang et al., 2016; Mariani and Borghi, 2021; Mariani et al., 2016; Mariani et al., 2018). The most popular form of BD from UGC is ORs, and the importance and function of these are examined in the next section.

Electronic word-of-mouth and online reviews in tourism and hospitality

The development and growth of the internet, as well as digital platforms and social media, have led to a widespread proliferation of UGC in the form of ORs. Online reviews enable consumers to develop, articulate and share their opinions about goods, services, firms and brands in online contexts (Hennig-Thurau et al., 2004). In marketing, ORs constitute an important part of eWOM, subsequently termed as online word-of-mouth (King et al., 2014).

Researchers in disciplines such as marketing, computer science and information management have analyzed both drivers and outcomes of eWOM (Hennig-Thurau et al., 2004; Rosario et al., 2016), and eWOM is much more dominant than simple WOM due to its rapidity of diffusion, potential anonymity, one-to-many and many-to-many reach, convenience, lack of face-to-face interaction, and communication effectiveness (Sun et al., 2006).

In the travel and tourism domains, eWOM is endlessly engendered through ORs produced by consumers on digital platforms, including online travel agencies (OTAs) such as Booking.com and Expedia.com, and online travel review websites such as Tripadvisor and Ctrip. Online reviews have been found to influence consumers' reservation intentions, sales (Ye et al., 2009) and firms' performance (Mariani and Borghi, 2020; Yang et al., 2018). Tourists and travelers deploy eWOM in the guise of ORs before purchase. For instance, it has been found that Tripadvisor ORs are commonly adopted by consumers to inform their accommodation decisions (Gretzel & Yoo, 2008) and generally that some of the features of ORs (such as ratings and volume) have a positive impact on commercial and financial performance (Yang et al., 2018).

Environmental concerns in travel, tourism and hospitality

Environmental awareness and perceptions in tourists' and travelers' electronic word-of-mouth

Over the last four decades an increasing number of firms, NGOs and national and international governments have adopted environmental schemes and practices (Pedersen & Neergaard, 2006) in response to the consolidation of an ecological position within the mainstream academic community recognizing that human consumption activity is a primary cause of natural resources depletion (United Nations, 2013) that might jeopardize not only firm growth and sustainability (Camilleri, 2018; Ek Styvén and Mariani, 2020; Salimath & Chandna, 2021), but also the very same existence of the planet.

Several policy makers, industry and professional associations, and other stakeholders (Clarkson, 1995) active in the tourism domain, realize that the consumption of natural resources is out of control, with negative repercussions on the planet (Di Pietro et al., 2013; Hardy et al., 2002). Accordingly, at the macro (national) level (Intergovernmental Panel on Climate Change, 2013), at the meso (organizational) level (Ettinger et al., 2018) and at the micro (individual) level (Liu et al., 2013) organizations and individuals are environmentally concerned. More specifically,

policy makers are trying to put pressure on businesses to embrace low environmental business practices (Gössling et al., 2016).

Firms are not only embracing more sustainable business processes and practices (Bonilla Priego et al., 2011) but increasingly showcasing what they do for the environment through CSR reporting (Guix et al., 2018; Thongplew et al., 2017; Watts, 2015) and by offering green products (Kemper et al., 2019). Indeed, it seems that industry discourse on environmental issues has, at least formally, been shifting towards a more serious commitment (Jones et al., 2014) than a decade ago when companies tended to downplay their environmental footprint (Gössling & Peeters, 2007). This is becoming increasingly relevant for both large hotel chains and SMEs in hospitality (Dief & Font, 2010). For instance, in addition to identifying four strategic priorities for its Planet 21 program (encompassing employees, customers, partners and local communities), the Accor hotel group has developed a strategic framework, including sustainable food and sustainable buildings. Marriott developed its “Make a Green Choice” initiative, allowing guests to give up housekeeping services (e.g. the use of cleaning products) and barter it for loyalty points or food and beverage vouchers. Also, independent hotels are becoming more active from an environmental stance. For example, independent hotels affiliated with the Greater Miami and the Beaches Hotel Association have launched sustainable practices aimed at reducing water consumption and water wasting products/processes (Raub & Martin-Rios, 2019).

Consumers are heterogeneous in the way they perceive and deal with environmental issues at a global level. Some are not entirely aware of environmental ecological issues or the relevance of environmental and sustainability practices and certifications (Di Pietro et al., 2013; Guizzardi et al., 2017); others are particularly sensitive to environmentally friendly tourism products and services (Dolnicar et al., 2008). Several studies in the tourism and hospitality domain (Di Pietro et al., 2013; Gustin & Weaver, 1996; Kyung et al., 2012; Masau & Prideaux, 2003) indicate that environmental concerns influence consumers’ purchase intentions. Several tourists are willing to pay a premium for green products and to meet sustainability requirements (Masau & Prideaux, 2003), whereas others think that the cost of sustainability initiatives should be incurred entirely by product/service providers (Gustin & Weaver, 1996). For instance, in a survey of U.S. hotel customers, Gustin and Weaver (1996) detected a positive relationship between customers’ attitudes towards U.S. hotels’ pro-environmental practices and purchase intentions. Kyung et al. (2012) find that, by deploying the New Ecological Paradigm (NEP) scale (Dunlap et al., 2000) on a sample of 9,000 tourists (the final respondents retained being 455) in Arizona, Florida and Texas, the degree of environmental concern positively influences the disposition to pay a premium for a hotel’s green initiatives, consistent with previous findings (Dutta et al., 2008; Gustin & Weaver, 1996). In their study of restaurant customers in the southeastern United States, Di Pietro et al. (2013) find that customers perceived that they were sufficiently knowledgeable about green practices, but they wanted to gain more knowledge about them. The respondents mentioned that they were more prone to choose restaurants that are environmentally friendly and use environmentally safe products. Environmentally aware customers were characterized as having higher levels of education and adopting green practices at home. Interestingly, even among consumers sensitive to environmentally friendly tourism products and services, only a tiny fraction of consumers are willing to pay for environmentally friendly products and services (Kyung et al., 2012).

Overall consumer research in tourism and hospitality has yielded mixed findings; several studies have emphasized that consumers are not aware of/knowledgeable about environmental issues and practices (e.g. Guizzardi et al., 2017; Prayag et al., 2022), whereas other investigations have pointed out that tourism and travelers are sensitive to environmentally friendly tourism products and services (e.g. Dolnicar et al., 2008; Kyung et al., 2012). This research puzzle seems to be the result of a confluence of several factors: (1) most of the studies carried out so far have focused on stated intentions and stated behaviors – they focused on consumers’ perceptions rather than behaviors (Pedersen & Neergaard, 2006); (2) the majority of the research produced

has examined perceptions *before* the service consumption (e.g. Di Pietro et al., 2013); (3) when Ors have been used as a source to document hotel guests' comments on service providers' CSR and environmental practices, convenience samples of customers have been used, adopting single case studies focusing on a few companies or destination cities within a specific continent (Brazytè et al., 2017; D'Acunto et al., 2020; Ettinger et al., 2018; Lee et al., 2016; Yu et al., 2017). For instance, Lee et al. (2016) analyze the top 10 green hotels in the U.S.A. on Tripadvisor and find that the majority of hotel guests respond in a positive way to green practices (such as saving energy and water) when they recognize them. Nonetheless, they can express negative opinions when they are not made aware of green practices, and this happens because not all the 10 hotels effectively inform prospective customers of green practices on Tripadvisor.

Online environmental discourse of tourists and travelers, and online review ratings

Research on UGC and, more specifically, Ors has examined the relationship between several characteristics and features of online reviews (and online reviewers) and online review scores as a proxy for online customer satisfaction with, and evaluation of, travel, tourism and hospitality services (e.g. Guo et al., 2017; Li et al., 2013; Radojevic et al., 2015; Xiang et al., 2017) and subsequently, looks at the capability of these to influence tourism and hospitality firm performance more than traditional (offline) customer satisfaction (Kim & Park, 2017).

So far, OR ratings have been mostly used to capture, assess and understand online consumer satisfaction with a specific travel, tourism and hospitality service or its constituent attributes (e.g. Li et al., 2013; Radojevic et al., 2015). Very few studies have tried to relate the wider CSR construct (including both the social and environmental dimensions) to online ratings (Brazytè et al., 2017; D'Acunto et al., 2020; Peiró-Signes et al., 2014). Leveraging OR ratings related to Spanish hotels, Peiró-Signes et al. (2014) find that four-star hotels endowed with an ISO 14001 environmental certification display higher ratings than those not endowed with such a certification. The difference is not significant in three- and five-star hotels, thus leading the authors to conclude that only four-star hotels are able to secure a distinctive competitive advantage from environmental certification.

By analyzing a small sample of 727 green ORs (i.e. ORs mentioning green experiences) covering the top 10 green hotels on Tripadvisor, Yu et al. (2017) mainly conduct descriptive and limited regression analyses and find that consumers document both positive and negative experiences at green hotels, and different types of green practices affect online customer satisfaction differently. More specifically, "energy" and "education and innovation" positively affect customer satisfaction, while "guest training," "energy," "water" and "purchasing" negatively influence customer satisfaction.

Building on a relatively small sample of less than 2,500 ORs of 30 Costa-Rican hotels endowed with a sustainable tourism certification, Brazytè et al. (2017) observe that less than a third mention implicitly sustainability-related indicators. They also find that ORs explicitly mentioning sustainability indicators, display higher ratings (average rating = 4.75) than reviews where sustainability indicators are not mentioned or implicitly mentioned (average rating = 4.39). The results seem to suggest that service customers that are aware of the sustainability actions taken by hotels are comparatively more satisfied than those where hotels/reviews did not mention, or only implicitly mentioned sustainability issues.

Based on content analysis of about 1,380 ORs related to 47 Austrian hotels emphasizing CSR aspects, Ettinger et al. (2018) find that the bulk of ORs – more than 90% – display a neutral/positive characterization. Yet, the authors do not examine what neutral/positive generates in terms of actual ratings.

Overall, research on the relationship between environmental eWOM and OR ratings is rather fragmented and limited. Indeed, it has been studied either in a specific country (e.g. Yu et al., 2017) or in a limited set of cities, within the same continent and using a limited number of

reviews compared to those available. Furthermore, as far as analytical techniques are concerned, the analyses conducted are extremely rudimentary (mostly relying on simple correlations, such as D'Acunto et al., 2020, Peiró-Signes et al., 2014) and when a model was fitted (Yu et al., 2017), it included a very small sample of observations.

Methodology

Data and sample

Online review data was collected from two distinctively different platforms: Booking.com and Tripadvisor. Both are good exemplars of a transaction-based and a community-based digital platform, respectively (Gligorijevic, 2016). The former was chosen because it hosts the highest share of certified hotel reviews worldwide (Revinat, 2018), while the latter was selected because it is the major online travel review website in the world. Data was collected through two scrapers developed in the Python programming language (by leveraging the libraries Selenium and BeautifulSoup) at the beginning of 2019. First, we sampled the top 10 city tourism destinations in terms of international tourist arrivals in both America and Europe (Geerts, 2018), and later focused on four city destinations in each: Barcelona, London, Paris and Rome (Europe); Las Vegas, Miami, New York City and Orlando (North America). We used the crawlers to retrieve the full list of reviewed hotels across our reference platforms.

Second, we collected the entire population of ORs covering the hotels located in these destinations across both platforms (i.e. Booking.com and Tripadvisor), over two years (2017–2018). Third, in line with other studies adopting text analytics (e.g. Xiang et al., 2017; Zhao et al., 2019), we kept only ORs written in English in our final database. We performed this task adopting the language detection package (langdetect) available in Python, which allows the detection of the language of a review through a lexicon-based analysis (Xiang et al., 2017). Overall, 2,702,227 ORs were retrieved: 1,144,461 from Tripadvisor and 1,557,766 from Booking.com. Therefore, the data builds on a very large number of reviews extracted from different types of OR platforms (community-based vs. transaction-based OR platform) and pertaining to hospitality services across multiple countries and continents.

Techniques adopted

To address our research question, we deployed model specifications with variables that are illustrated in the following section and, more specifically, Tobit regression analysis (Tobin, 1958) for the Booking.com ORs (as the dependent variable of rating is close to continuous but is left and right censored) and ordinal logistic regression analysis (Greene, 1999) for the Tripadvisor ORs (as the dependent variable is ordinal and can assume only five categorical values).

Variables

The key variables used in this study are illustrated and described in Table 1. The two focal variables – *Environmental Presence* and *Environmental Depth* – are operationalized in line with extant literature (Mariani & Borghi, 2021a). Following the lead of Gao et al. (2018), the *Observed Average Rating* is defined as the average hotel's OR rating, as observed by the reviewer at the time when they posted their review. *Reviewer Experience* is a proxy of the experience in online reviewing in the focal OR platform and it is measured as the overall number of ORs written by the reviewer in the focal platform (i.e. either Tripadvisor or Booking.com). *Reviewer Image* is a dummy variable that equals to 1 if the reviewer used a personalized image for their profile on the platform

Table 1. Variables and descriptions.

Variable	Description
Environmental Presence	It is a dummy variable that is equal to 1 if the review includes at least one word in the environmental dictionary developed by Pencle and Mălăescu (2016), and zero otherwise (see Mariani and Borghi, 2020).
Environmental Depth	It is a ratio equal to the number of environment-related words (words present in the Pencle & Mălăescu, 2016 environmental dictionary) over the total number of words in the review, multiplied by 100. In other words, it captures the percentage of environmental elements in the online review (see Mariani and Borghi, 2020).
Observed Average Rating (Observed Avg Rating)	Hotels' review average rating as observed by the reviewing guest at the time when they posted their review (see Gao et al., 2018)
Reviewer Experience	It is the overall number of online reviews written by the reviewer in the platform.
Reviewer Image	It is a dummy variable that is equal to 1 if the reviewer used a personalized image for its social profile in the platform, and zero otherwise (Forman et al., 2008).
Country of origin disclosure (Country Disclosure)	It is a dummy variable that is equal to 1 if the reviewers did not disclose their country of origin, and zero otherwise (see Filieri et al., 2019).
Review Length	It represents the number of words included in each online review (Chevalier & Mayzlin, 2006; Zhang et al., 2016).
Review Polarity	The <i>polarity</i> , also known as sentiment score, was operationalized using a continuous variable ranging from -1 to $+1$, respectively, equating to extremely negative and extremely positive content and emotions. To create this measure, we used the Valence aware dictionary for sentiment reasoning (VADER), which exploits a set of heuristics along with a specific lexicon dictionary for this particular task (Hutto & Gilbert, 2014).
Type of Trip	It is a categorical variables which embeds two options: Leisure and Business
Type of Group	It is a categorical variable which embeds 4 different options: Couple, Solo, Family and Group.
Submission Device	It is a dummy variable that is equal to 1 if the review has been written using a mobile device, and zero otherwise
Destination City	It is a categorical variable that indicates the city where the reviewed hotel is located.
Year	It represents the year in which the review has been written.
Chain	A dummy variable that is equal to 1 if the hotel is part of a chain company and zero otherwise.
Star Rating	It is a categorical variable that describes the hotel class category adopted to classify hotels according to their quality (from 1- to 5-stars). It has been retrieved directly from the page of each hotel as displayed in the online review platform (Booking.com or TripAdvisor).a
Review Valence	Online review rating posted by online reviewers to summarize with a number their satisfaction with the hospitality service.

^aSince TripAdvisor uses a customized rating system including half-star rating, in order to obtain the final value we rounded the retrieved value to the closest half-star rating. (Gao et al., 2018).

(Forman et al., 2008). *Country of Origin Disclosure* is a dummy variable that is equal to 1 if the reviewer did not disclose their country of origin (see Filieri et al., 2019).

In terms of text analytics, we focused on *Review Length* and *Review Polarity*. The former consists of the number of words in the OR (e.g. Chevalier & Mayzlin, 2006; Zhang et al., 2016). The latter, sometimes termed as sentiment score, is a continuous variable ranging from -1 to $+1$ and was computed using the Valence Aware Dictionary for sEntiment Reasoning (VADER), based on a dictionary and heuristics (Hutto & Gilbert, 2014). The measure was preferred over other alternative measures as recent research has found that it performs other measures in the tourism domain (Alaei et al., 2019).

Several control variables were used, including *Type of Trip*, *Type of Group*, *Submission Device*, *Destination City*, *Year*, *Chain* and *Star Rating*. *Type of Trip* is a categorical variable that describes the trip purpose: leisure or business. *Type of Group* is a categorical variable that indicates if the travelers were a couple, solo, family or group. The *Submission Device* is a dummy variable that describes if the OR was written using a mobile device or a different device connected to the

Table 2a. Descriptive statistics for the Tripadvisor sample, 2017–2018.

	Mean	Median	SD	Min	Max
Valence	4.120	5.000	1.164	1	5
Environmental Depth	0.998	0.550	1.307	0	19.350
Log Environmental Depth	0.522	0.438	0.560	0	3.013
Environmental Presence	0.538	1	0.499	0	1
Observed Average Rating	4.123	4.217	0.526	1	5
Reviewer Expertise	39.707	9.000	116.219	1	10756.000
Log Reviewer Expertise	2.204	2.197	1.747	0	9.283
Reviewer Image	0.278	0	0.448	0	1
Country Disclosure	0.766	1	0.423	0	1
Review Length	108.103	73.000	103.693	1	3603.000
Log Review Length	4.411	4.290	0.684	0	8.190
Review Polarity	0.742	0.946	0.484	-0.999	1
Observations	1,144,461				

Table 2b. Descriptive statistics for the Booking.com sample, 2017–2018.

	Mean	Median	SD	Min	Max
Valence	7.950	8.300	1.919	2.5	10
Environmental Depth	1.047	0	3.110	0	100.000
Log Environmental Depth	0.352	0	0.687	0	4.615
Environmental Presence	0.248	0	0.432	0	1
Observed Average Rating	7.954	8.137	0.942	2.5	10
Reviewer Expertise	7.617	3.000	12.433	1	1047.000
Log Reviewer Expertise	1.330	1.099	1.136	0	6.954
Reviewer Image	0.411	0	0.492	0	1
Country Disclosure	0.999	1	0.035	0	1
Review Length	34.454	21.000	41.340	1	772.000
Log Review Length	2.956	3.045	1.157	0	6.649
Review Polarity	0.392	0.557	0.516	-0.998	0.999
Observations	1,557,766				

internet (e.g. desktop computer). *Destination City* is a categorical variable that indicates the city where the reviewed hotel is located. *Year* represents the year when the review was written. *Chain* is a dummy variable that indicates whether the hotel belongs to a hotel chain or not. *Star Rating* is a categorical variable that describes the hotel class category adopted to classify hotels according to their quality (from one to five stars). Overall, the controls have been used in extant tourism and hospitality literature trying to identify and explain the determinants of online review ratings (e.g. Gao et al., 2018; Mariani et al., 2019a,2019b)

Online Review Valence has been used as a dependent variable to address our research question and it represents the OR rating posted by an online reviewer to summarize with a number their satisfaction with the hospitality service.

Finally, environmental depth, reviewer experience and review length were log transformed, given the skewness of the variables’ distribution (Filieri et al., 2019). Tables 2.a and 2.b illustrate the descriptive statistics of the variables under consideration for the period 2017–2018.

Model specification

When choosing the econometric model to deploy, we considered the nature of the dependent variable analyzed in the study, namely the overall online rating. Thus, on the one hand, when analyzing Booking.com ORs, we adopted a Tobit multivariate regression as the online rating is continuous but both left and right censored, with the minimum and maximum variables being 2.5 and 10.0, respectively (Mariani & Borghi, 2018). On the other hand, for Tripadvisor ORs we used an ordered logit regression model as the online rating in the platform is a categorical

variable assuming only five values (see Yu et al., 2017; Zhang et al., 2016). Therefore, based on the samples indicated in the research design, we developed two model specifications.

More formally, denoting with $Valence_{ij}^*$ the underlying latent variable representing the latent overall rating provided by reviewer i for hotel j , we estimated the following econometric specifications:

$$\begin{aligned}
 Valence_{ij}^* = & \beta_0 + \beta_1(Environmental\ Presence)_{ij} + \beta_2(Observed\ Avg\ rating)_{ij} \\
 & + \beta_3(Reviewer\ Experience)_{ij} + \beta_4(Reviewer\ Image)_{ij} + \beta_5(Country\ Disclosure)_{ij} \\
 & + \beta_6(Review\ Length)_{ij} + \beta_7(Review\ Polarity)_{ij} \\
 & + \beta_8(Submission\ Device)_{ij} \\
 & + \beta_9(Year\ Dummy)_{ij} + \beta_{10}(Chain)_j + \theta'_1(Type\ of\ Trip)_{ij} + \theta'_2(Type\ of\ Group)_{ij} \\
 & + \theta'_3(Destination\ City)_j + \theta'_4(Star\ Rating)_j + \epsilon_{ij}
 \end{aligned} \tag{Eq. 1}$$

$$\begin{aligned}
 Valence_{ij}^* = & \beta_0 + \beta_1(Environmental\ Depth)_{ij} + \beta_2(Observed\ Avg\ rating)_{ij} \\
 & + \beta_3(Reviewer\ Experience)_{ij} + \beta_4(Reviewer\ Image)_{ij} + \beta_5(Country\ Disclosure)_{ij} \\
 & + \beta_6(Review\ Length)_{ij} + \beta_7(Review\ Polarity)_{ij} \\
 & + \beta_8(Submission\ Device)_{ij} \\
 & + \beta_9(Year\ Dummy)_{ij} + \beta_{10}(Chain)_j + \theta'_1(Type\ of\ Trip)_{ij} + \theta'_2(Type\ of\ Group)_{ij} \\
 & + \theta'_3(Destination\ City)_j + \theta'_4(Star\ Rating)_j + \epsilon_{ij}
 \end{aligned} \tag{Eq. 2}$$

where ϵ_{ij} refers to the error term at the single review level and, β_i and θ_i indicate the regression coefficients and vector of coefficients, respectively. From the latent variable $Valence_{ij}^*$, in the case of the ordered logit regression model, the observed review rating $Valence_{ij}$, is obtained through the following rule:

$$Valence_{ij} = r \text{ if } k_{r-1} < Valence_{ij}^* \leq k_r, \text{ for } r = 1, \dots, R$$

where r relates to the number of alternatives embedded in the model, in our case $R=5$ since Tripadvisor allows to rate a hotel on a five-point Likert scale. Besides, we assume $k_0 = -\infty$ and $k_R = \infty$, whereas k_1 to k_{R-1} designate the set of cut-off thresholds, estimated by the model, that are used to determine the correct interval of materialization of the predicted latent variable (Cameron & Trivedi, 2005).

Regarding the Tobit model, the observed Booking.com rating for an online review is specified as:

$$Valence_{ij} = \begin{cases} Valence_{ij}^* & \text{if } y_L < Valence_{ij}^* < y_U \\ y_L & \text{if } Valence_{ij}^* \leq y_L \\ y_U & \text{if } Valence_{ij}^* \geq y_U \end{cases}$$

where y_L and y_U represent the lower (2.5) and upper (10) limit of the distribution of the observed dependent variable, respectively.

As clear from both models, the reference dependent variable is online review valence (namely the OR ratings) and it was regressed against the focal independent variables (Environmental Discourse Presence in model 1 and Environmental Discourse Presence in model 2) and a series of other explanatory variables, including observed average rating, reviewer experience, reviewer image, country disclosure, text analytics (namely Review Length and Review Polarity) and control variables such as type of trip, type of group, submission device, destination city, year, chain and star rating.

Table 3 Online review ratings and environmental discourse depth and presence in Tripadvisor and Booking.com eWOM, 2017–2018.

	Tripadvisor		Booking.com	
	(1.T)	(2.T)	(1.B)	(2.B)
Environmental Presence	0.132*** (0.00417)		0.0186*** (0.00359)	
Log Environmental Depth		0.131*** (0.00355)		0.0277*** (0.00212)
Observed Avg Rating	1.417*** (0.00434)	1.416*** (0.00434)	0.833*** (0.00184)	0.832*** (0.00184)
Log Reviewer Experience	-0.0757*** (0.00133)	-0.0760*** (0.00133)	-0.00950*** (0.00129)	-0.00958*** (0.00129)
Reviewer Image	0.0526*** (0.00464)	0.0526*** (0.00464)	0.0636*** (0.00290)	0.0637*** (0.00290)
Country Disclosure	0.0296*** (0.00493)	0.0299*** (0.00493)	0.143*** (0.0407)	0.143*** (0.0407)
Log Review Length	-0.737*** (0.00320)	-0.708*** (0.00301)	-0.404*** (0.00138)	-0.405*** (0.00130)
Review Polarity	2.550*** (0.00491)	2.548*** (0.00491)	1.955*** (0.00292)	1.954*** (0.00293)
Control Variables				
Type of Trip	YES	YES	YES	YES
Type of Group	YES	YES	YES	YES
Submission Device	YES	YES	YES	YES
Destination Country	YES	YES	YES	YES
Year	YES	YES	YES	YES
Chain	YES	YES	YES	YES
Star Rating	YES	YES	YES	YES
Constant			1.826*** (0.0443)	1.828*** (0.0443)
Intercept-1	0.0401 (0.0274)	0.158*** (0.0272)		
Intercept-2	1.325*** (0.0273)	1.443*** (0.0271)		
Intercept-3	2.864*** (0.0274)	2.982*** (0.0273)		
Intercept-4	4.655*** (0.0277)	4.774*** (0.0275)		
Observations	1,144,461	1,144,461	1,557,766	1,557,766
Pseudo R ²	22.7%	22.7%	13.3%	13.3%
AIC	2,232,788.1	2,232,429.7	5,407,878.6	5,408,021.9
LR Chi2	655,268.1	655,626.5	829,706.5	829,563.2
Log Likelihood	-1,116,363.0	-1,116,183.8	-2,703,911.3	-2,703,982.9

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Key findings

The results of the regression analyses capturing the influence of environmental discourse *presence* and environmental discourse *depth* on OR ratings are illustrated in Table 3. Online consumers' environmental discourse *presence* appears to influence positively OR ratings, both on Booking.com ($p < 0.001$) and on Tripadvisor ($p < 0.001$) (see, respectively, models 1.B and 1.T in Table 3). This finding extends extant small-sample research that found that environmental certifications or wider CSR practices positively influence OR ratings (e.g. Brazytė et al., 2017; Peiró-Signes et al., 2014).

Interestingly, online consumers' environmental discourse *presence* appears to influence positively OR ratings both on Booking.com ($p < 0.001$) and on Tripadvisor ($p < 0.001$) (see models 2.B and 2.T in Table 3). This finding is rather novel if benchmarked with extant literature in the tourism and hospitality field and suggests that the comprehensiveness of the environmental-related discourse within ORs has the capability to influence positively and significantly online review ratings.

The analysis of the focal independent variables (Environmental Presence and Depth) suggests that both the *presence* (proxying environmental awareness) and the *depth* (proxying in-depth digging about environmental aspects) of environmental-related aspects in ORs make a difference for online ratings and, ultimately, for customers' online satisfaction. The effects measured are robust regardless of the digital platform considered (community vs. transaction-based platform), geographic setting (continent or country where the destination is located), type of trip (leisure or business), type of traveling group (couple, solo, family and group), submission device (mobile vs. desktop), time (specific year), or firm level characteristics (such as belonging to a chain or to a specific category).

In relation to the reviewer-level control variables, reviewers' experience in online reviewing is negative and significant ($p < 0.001$) in all the four models, consistent with previous research (e.g. Mariani et al., 2019b). Given that experience in online reviewing is, to a certain extent, connected to travel experience (Gao et al., 2018), this result is consistent with extant research showing that experience has a negative impact on reviewers' online ratings.

A reviewer's disclosure of their image can positively influence online ratings ($p < 0.001$) in all the four models, in line with extant research conducted in tourism and hospitality settings (Gao et al., 2018). Disclosure of the reviewer's country of origin also exerts a positive impact on online ratings ($p < 0.001$) in each of the four models, consistent with the existing body of tourism management and marketing literature (Fileri et al., 2019).

As far as the text analytics are concerned, *Review Length* negatively influences ($p < 0.001$) OR ratings in all the four models, and this is considered in line with extant literature (Chevalier & Mayzlin, 2006) as customers tend to put more effort in writing and write longer reviews when they are dissatisfied with a product or service. *Review Polarity* positively affects ($p < 0.001$) OR ratings in all the four models, again considered in line with extant literature (Geetha et al., 2017) as customers evaluate their consumption experience more positively when they are in a positive emotional state (Isen, 1987).

Discussion and conclusion

Building on more than 2.7 million ORs collected from Booking.com and Tripadvisor.com (Li et al., 2018; Mariani et al., 2018), and examined with sophisticated big data analytical techniques, this work has described consumers' behaviors and how consumers perceive and deal with sustainability issues, and if consumers' environmental discourse affects their OR ratings. By adopting multivariate regression analyses, we have shown that OR ratings are positively influenced by both environmental discourse *presence* and *depth* on both the platforms analyzed. Overall, our findings shed light on the *presence* and impact on customer satisfaction of consumers' environmental concerns embedded in tourists' UGC, across several leading tourism destinations based in two different continents.

Theoretical contributions

This study contributes to digital platforms, BDA, eWOM and environmental research within the hospitality and tourism domain in multiple ways. To the best of our knowledge, this is one of the first studies in sustainable tourism research addressing how the *presence* and *depth* of online consumers' environmental discourse emerge on digital platforms, thus complementing a nascent research line (Mariani & Borghi, 2021a) which simply tracked longitudinally the *presence* and *depth* of online consumers' environmental discourse. Unlike previous studies that have used social media (predominantly Twitter) to track online customers' opinions (e.g. Chisholm & O'Sullivan, 2017; Reyes-Menendez et al., 2018), this is a cross-platform study ensuring robustness of findings and generalizability. Moreover, while previous studies (e.g. Mariani & Borghi, 2021a)

measured the presence and depth of the environmental discourse, the present work measures quantitatively the influence of the *presence* and *depth* of the environmental discourse on online review ratings, which constitute a proxy of customer satisfaction. Second, our findings seem to broadly suggest that tourists are interested in (and engage with) environmental issues and have environmental concerns, voicing them – either explicitly or implicitly – through their ORs; this finding resonates with recent literature emphasizing that travelers seek sustainable experiences (Oklevik et al., 2019). This extends recent studies (Mariani & Borghi, 2021a) that have looked at how digital technologies and digitalization can help monitor online consumers' environmental discourse, by suggesting that digital technologies – and especially BDA as a pillar of the digital transformation – can help consumers and tourism researchers make sense of online reviewers' evaluation of sustainable tourism products, services and experiences. Third, we indicate that in today's digital world, consumers' concerns can be captured effectively by relying on digital platforms and ecosystems (Nambisan, 2017) which offer a rich data source in the guise of consumer reviews. This seems to support earlier conceptualization of virtual customers (Nambisan & Baron, 2007; Nambisan & Nambisan, 2008) that have played a critical role in the development of digital entrepreneurship literature (Nambisan, 2017). Fourth, by focusing on the online tourists' environmental concerns, this work enriches the body of eWOM literature in the sustainable tourism field that, so far, has mainly focused on broad CSR discourses (e.g. Ettinger et al., 2018) or green practices at the local or, at best, national level (e.g. Lee et al., 2016; Yu et al., 2017). Building on the received distinction between the *presence* and the *depth* of online consumers' environmental discourse (Mariani & Borghi, 2021a), we test a model and offer empirical substantiation of the constructs/variables developed in the recent sustainable tourism literature. Fifth, we contribute to consumer behavior literature (e.g. Straughan & Roberts, 1999; Thøgersen et al., 2010) by suggesting that environmental issues do influence consumer perceptions of tourism and hospitality services. Sixth, we contribute to literature on big data and analytics in tourism (Li et al., 2018; Mariani et al., 2018), by methodologically developing and deploying text analytics from UGC sourced across different digital platforms. Seventh, we enrich the emerging debate on consumers' perception of CSR initiatives in the tourism and hospitality industry (Ettinger et al., 2018), especially the research stream at the intersection between online consumers' perceptions of environmental issues through eWOM and their satisfaction with hospitality and tourism services (e.g. Brazytè et al., 2017; Lee et al., 2016; Yu et al., 2017), by shedding new light on the relationship between environmental discourse presence and online customer satisfaction, proxied by means of OR ratings. In addition to augmenting the power and generalizability of previous findings on the relationship between green practices and OR hotel ratings (by adopting the largest sample of ORs considered so far in the research stream), we go much beyond and – based on well specified multivariate regression models – also evaluate the relationship between the *depth* of the environmental discourse embedded in ORs on online consumer review ratings. Text analytics suggest that both the *presence* and *depth* of online consumers' environmental discourse positively and consistently effect OR ratings.

Practical implications

Several practical implications stem from this work, including implications for tourism and hospitality practitioners, and digital platforms managers and developers.

As far as tourism and hospitality practitioners are concerned, they should recognize that travelers give better evaluations to services when environmental aspects are mentioned, as is clear from the positive and significant effect of environmental discourse presence and depth within online review ratings across both the platforms analyzed. This should push hotel managers to: 1) invest in and develop green initiatives, programs and practices, with the awareness that only a tiny portion of consumers will pay for eco-friendly products and services (Kyung et al., 2012); 2)

look for environmental certification, as this has been found to generate benefits in specific conditions (Peiró-Signes et al., 2014); 3) develop a clear marketing communication strategy and ad hoc communication tactics that encompass clear messages aimed at clarifying and detailing prospective customers' tangible benefits pertaining to the green practices adopted (Dolnicar et al., 2017), beyond explaining the societal benefits of those green practices; 4) possibly include education in the repertory of their green practices; environmental education has been found to play a crucial role in affecting online ratings (Yu et al., 2017); 5) emphasize their environmental commitment when responding to (positive) ORs that have mentioned environmental aspects – this might eventually translate into higher purchase intentions of prospective customers interested in environmentally friendly experiences; 6) record preferences towards specific green practices (e.g. towel reuse) into the company database to facilitate instantaneous knowledge about the customer and their pro-environmental preferences.

Platform developers and managers that deal with OTAs and online community travel review platforms would benefit from the findings of this study as they host ORs which relate to environmental issues. As environmental-related ORs display different features than non-environmental-related ORs (as is clear from previous research), platform managers might develop and/or use environmental dictionaries, allowing them to segment both ORs and online reviewers that wrote them, as this might affect online review ratings. Second, given that the presence of environmental-related ORs seem to have a positive impact on online ratings, platforms developers can develop an attribute named “green friendly” for the hospitality service, as other research conducted in offline contexts has shown that the presence of a “green” attribute enhances consumers' involvement in the purchase process (Straughan & Roberts, 1999; Thøgersen et al., 2012). This might assist both hotel managers and customers in designing and assessing green processes and practices, respectively. Eventually, this will increase consumer bookings at hotels that explicitly cater to consumers with ecological sensitivities.

Conclusion and limitations

This work contributes insights into online consumers' perceptions of environmental concerns and sheds light on the extent to which consumers' environmental sensitivity, as expressed on digital platforms, can influence customer satisfaction. In particular, we measure empirically the effect of online consumers' environmental discourse on online review ratings as a proxy of customer satisfaction. The objective is pursued by leveraging large volumes (big data) of OR sources in relation to hospitality services across different firms, destinations, continents and countries. Accordingly, this paper contributes to the area at the intersection between digital platforms, BDA, eWOM and environmental research within the hospitality and tourism domain, by examining the relationship between the environmental features of ORs and OR ratings (Peiró-Signes et al., 2014; Yu et al., 2017).

Leveraging more than 2.7 million ORs collected and examined by means of BDA techniques, we detected that both the *presence* and *depth* of online consumers' environmental discourse have a positive impact on OR ratings. Interestingly, OR ratings are positively influenced by both environmental discourse on *presence* and *depth* on both platforms, regardless of the type of platform analyzed.

This study is not without limitations. First, while we have used several controls, further variables (such as demographics) could be included in the analysis. However, on both Booking.com and Tripadvisor most of the reviewers' profiles lack demographic information and, when present, the demographic information is not necessarily reliable. This is the reason why in other popular and recent studies using online review analytics (e.g. Xiang et al., 2017; Zhao et al., 2019), demographics are never used. Second, while this work represents the first attempt to measure the impact of the *presence* and *depth* of online consumers' environmental discourse on online review

ratings across different destinations and continents, other destinations in the Asia-Pacific could also be considered. Third, a common limitation when using dictionaries is that researchers assume the words used are related to the specific context, wherein the word is produced or to which the word relates to with the higher probability; it might be possible that words in the environmental dictionary might also relate to the destination. However, as they are included in ORs of hotels it is realistic to assume, like other researchers have done (e.g. D'Acunto et al., 2020), that the words mostly relate to the hotel.

Acknowledgements

We would like to thank the editors and the anonymous reviewers for their helpful and constructive comments. We are grateful to one colleague for her advice on earlier versions of the paper and to Emma Lang and Kathryn Pilgrim for their assistance with professional copyediting and design.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Marcello Mariani  <http://orcid.org/0000-0002-7916-2576>

References

- Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., & Weaven, S. (2019). Market segmentation and travel choice prediction in Spa hotels through TripAdvisor's online reviews. *International Journal of Hospitality Management*, 80, 52–77. <https://doi.org/10.1016/j.ijhm.2019.01.003>
- Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: Capitalizing on Big Data. *Journal of Travel Research*, 58(2), 175–191. <https://doi.org/10.1177/0047287517747753>
- Bonilla Priego, M. J., Najera, J. J., & Font, X. (2011). Environmental management decision-making in certified hotels. *Journal of Sustainable Tourism*, 19(3), 361–381. <https://doi.org/10.1080/09669582.2010.530350>
- Brazytė, K., Weber, F., & Schaffner, D. (2017). Sustainability management of hotels: How do customers respond in online reviews? *Journal of Quality Assurance in Hospitality & Tourism*, 18(3), 282–307. <https://doi.org/10.1080/1528008X.2016.1230033>
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Camilleri, M. A. (2018). Closing the loop for resource efficiency, sustainable consumption and production: A critical review of the circular economy. *International Journal of Sustainable Development*, 1(1), 1–17. <https://doi.org/10.1504/IJSD.2018.10012310>
- Chang, Y.-C., Ku, C.-H., & Chen, C.-H. (2020). Using deep learning and visual analytics to explore hotel reviews and responses. *Tourism Management*, 80, 104129. <https://doi.org/10.1016/j.tourman.2020.104129>
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3), 345–354. <https://doi.org/10.1509/jmkr.43.3.345>
- Chisholm, E., & O'Sullivan, K. (2017). Using Twitter to explore (un)healthy housing: Learning from the #Characterbuildings campaign in New Zealand. *International Journal of Environmental Research and Public Health*, 14(11), 1424.
- Clarkson, M. E. (1995). A stakeholder framework for analyzing and evaluating corporate social performance. *The Academy of Management Review*, 20(1), 92–117. <https://doi.org/10.2307/258888>
- D'Acunto, D., Tuan, A., Dalli, D., Viglia, G., & Okumus, F. (2020). Do consumers care about CSR in their online reviews? An empirical analysis. *International Journal of Hospitality Management*, 85, 102342. <https://doi.org/10.1016/j.ijhm.2019.102342>
- Davenport, T. H. (2014). How strategists use “big data” to support internal business decisions, discovery and production. *Strategy & Leadership*, 42(4), 45–50.
- Di Pietro, R. B., Cao, Y., & Partlow, C. (2013). Green practices in upscale foodservice operations. *International Journal of Contemporary Hospitality Management*, 25(5), 779–796. <https://doi.org/10.1108/IJCHM-May-2012-0082>
- Dief, M. E., & Font, X. (2010). The determinants of hotels' marketing managers' green marketing behavior. *Journal of Sustainable Tourism*, 18(1), 157–174. <https://doi.org/10.1080/09669580903464232>

- Dolnicar, S., Crouch, G. I., & Long, P. (2008). Environment-friendly tourists: What do we really know about them? *Journal of Sustainable Tourism*, 16(2), 197–210. <https://doi.org/10.2167/jost738.0>
- Dolnicar, S., Knezevic Cvelbar, L., & Grün, B. (2017). Do pro-environmental appeals trigger pro-environmental behavior in hotel guests? *Journal of Travel Research*, 56(8), 988–997. <https://doi.org/10.1177/0047287516678089>
- Dunlap, R. E., Van Liere, K. D., Mertig, A. G., & Jones, R. E. (2000). Measuring endorsement of the new ecological paradigm: A revised NEP scale. *Journal of Social Issues*, 56(3), 425–442. <https://doi.org/10.1111/0022-4537.00176>
- Dutta, K., Umashankar, V., Choi, G., & Parsa, H. G. (2008). A comparative study of consumers' green practice orientation in India and the United States: A study from the restaurant industry. *Journal of Foodservice Business Research*, 11(3), 269–285. <https://doi.org/10.1080/15378020802316570>
- Ek Styvén, M., & Mariani, M. M. (2020). Understanding the intention to buy secondhand clothing on sharing economy platforms: The influence of sustainability, distance from the consumption system, and economic motivations. *Psychology & Marketing*, 37(5), 724–739. <https://doi.org/10.1002/mar.21334>
- Ervellas, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897–904. <https://doi.org/10.1016/j.jbusres.2015.07.001>
- Ettinger, A., Grabner-Kräuter, S., & Terlutter, R. (2018). Online CSR communication in the hotel industry: Evidence from small hotels. *International Journal of Hospitality Management*, 68, 94–104. <https://doi.org/10.1016/j.ijhm.2017.09.002>
- Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, 52, 498–506. <https://doi.org/10.1016/j.tourman.2015.07.018>
- Filieri, R., Raguseo, E., & Vitari, C. (2019). What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *International Journal of Hospitality Management*, 77, 333–341. <https://doi.org/10.1016/j.ijhm.2018.07.013>
- Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, 19(3), 291–313. <https://doi.org/10.1287/isre.1080.0193>
- Fosso Wamba, S., Dubey, R., Gunasekaran, A., & Akter, S. (2020). The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism. *International Journal of Production Economics*, 222, 107498. <https://doi.org/10.1016/j.ijpe.2019.09.019>
- France, S. L., & Ghose, S. (2019). Marketing analytics: Methods, practice, implementation, and links to other fields. *Expert Systems with Applications*, 119, 456–475.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Gao, B., Li, X., Liu, S., & Fang, D. (2018). How power distance affects online hotel ratings: The positive moderating roles of hotel chain and reviewers' travel experience. *Tourism Management*, 65, 176–186. <https://doi.org/10.1016/j.tourman.2017.10.007>
- Geerts, W. (2018). Top 100 City Destinations 2018, Euromonitor International. Retrieved 11 January 2022, from <https://go.euromonitor.com/white-paper-travel-2018-100-cities>
- Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels – An empirical analysis. *Tourism Management*, 61, 43–54. <https://doi.org/10.1016/j.tourman.2016.12.022>
- Gligorijevic, B. (2016). Review platforms in destinations and hospitality. In R. Egger, I. Gula & D. Walcher, (Eds.), *Open Tourism*. (pp. 215–228). Springer.
- Gong, L., Liu, X., Wu, L., & Liu, Y. (2016). Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartography and Geographic Information Science*, 43(2), 103–114. <https://doi.org/10.1080/15230406.2015.1014424>
- Gorgemans, S., & Murillo-Luna, J. L. (2016). The adoption of environmental management systems by shelters. In M. Mariani, W. Czakon, D. Buhalis & O. Vitouladiti (Eds.), *Tourism Management, Marketing, and Development: Performance, Strategies, and Sustainability*. (pp. 259–275). Palgrave Macmillan.
- Gössling, S. (2017). Tourism, information technologies and sustainability: An exploratory review. *Journal of Sustainable Tourism*, 25(7), 1024–1041. <https://doi.org/10.1080/09669582.2015.1122017>
- Gössling, S., & Peeters, P. (2007). It does not harm the environment! An analysis of industry discourses on tourism, air travel and the environment. *Journal of Sustainable Tourism*, 15(4), 402–417. <https://doi.org/10.2167/jost672.0>
- Gössling, S., & Peeters, P. (2015). Assessing tourism's global environmental impact 1900–2050. *Journal of Sustainable Tourism*, 23(5), 639–659. <https://doi.org/10.1080/09669582.2015.1008500>
- Gössling, S., Cohen, S. A., & Hares, A. (2016). Inside the black box: EU policy officers' perspectives on transport and climate change mitigation. *Journal of Transport Geography*, 57, 83–93. <https://doi.org/10.1016/j.jtrangeo.2016.10.002>
- Greene, W. H. (1999). *Econometric Analysis*, 4th Edn. Prentice Hall.
- Gretzel, U., & Yoo, K. H. (2008). Use and impact of online travel reviews. *Proceedings of the International Information and Communication Technologies in Tourism Conference*, 35–46.

- Guix, M., Bonilla-Priego, M. J., & Font, X. (2018). The process of sustainability reporting in international hotel groups: An analysis of stakeholder inclusiveness, materiality and responsiveness. *Journal of Sustainable Tourism*, 26(7), 1063–1022. <https://doi.org/10.1080/09669582.2017.1410164>
- Guizzardi, A., Mariani, M. M., & Prayag, G. (2017). Environmental impacts and certification: Evidence from the Milan World Expo 2015. *International Journal of Contemporary Hospitality Management*, 29(3), 1052–1071. <https://doi.org/10.1108/IJCHM-09-2015-0491>
- Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, 59, 467–483. <https://doi.org/10.1016/j.tourman.2016.09.009>
- Gustin, M. E., & Weaver, P. A. (1996). Are hotels prepared for the environmental consumer? *Hospitality Research Journal*, 20(2), 1–14. <https://doi.org/10.1177/109634809602000201>
- Hao, J.-X., Fu, Y., Hsu, C., Li, X., & Chen, N. (2020). Introducing news media sentiment analytics to residents' attitudes research. *Journal of Travel Research*, 59(8), 1353–1369. <https://doi.org/10.1177/0047287519884657>
- Hardy, A., Beeton, R. J., & Pearson, L. (2002). Sustainable tourism: An overview of the concept and its position in relation to conceptualisations of tourism. *Journal of Sustainable Tourism*, 10(6), 475–496. <https://doi.org/10.1080/09669580208667183>
- Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, 18(1), 38–52. <https://doi.org/10.1002/dir.10073>
- Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. Retrieved 12 January 2022, from <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>
- Intergovernmental Panel on Climate Change (2013). *Climate Change 2013. The Physical Science Basis*. Cambridge University Press.
- Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 20, 203–253.
- Jones, P., Hillier, D., & Comfort, D. (2014). Sustainability in the global hotel industry. *International Journal of Contemporary Hospitality Management*, 26(1), 5–17. <https://doi.org/10.1108/IJCHM-10-2012-0180>
- Kasahara, H., Mori, M., Mukunoki, M., & Minoh, M. (2015). Transportation mode annotation of tourist GPS trajectories under environmental constraints. In A. Inversini, & R. Schegg (Eds.), *Information and Communication Technologies in Tourism 2015*. (pp. 523–535). Springer.
- Kemper, J. A., Hall, C. M., & Ballantine, P. W. (2019). Marketing and sustainability: Business as usual or changing worldviews? *Sustainability*, 11(3), 780. <https://doi.org/10.3390/su11030780>
- Kenney, M., & Zysman, J. (2016). The rise of the platform economy. *Issues in Science and Technology*, 32(3), 61–69.
- Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction: Which one has more incremental predictive power in explaining hotel performance? *International Journal of Contemporary Hospitality Management*, 29(2), 784–802. <https://doi.org/10.1108/IJCHM-11-2015-0627>
- King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing*, 28(3), 167–183. <https://doi.org/10.1016/j.intmar.2014.02.001>
- Kyung, H. K., Stein, L., Heo, C. Y., & Lee, S. (2012). Consumers' willingness to pay for green initiatives of the hotel industry. *International Journal of Hospitality Management*, 31(2), 564–572. <https://doi.org/10.1016/j.ijhm.2011.08.001>
- Lee, H., Jai, T. M., & Li, X. (2016). Guests' perceptions of green hotel practices and management responses on TripAdvisor. *Journal of Hospitality and Tourism Technology*, 7(2), 182–199. <https://doi.org/10.1108/JHTT-10-2015-0038>
- Li, H. Y., Ye, Q., & Law, R. (2013). Determinants of customer satisfaction in the hotel industry: An application of online review analysis. *Asia Pacific Journal of Tourism Research*, 181. (7), 784–802. <https://doi.org/10.1080/10941665.2012.708351>
- Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, 68, 301–323. <https://doi.org/10.1016/j.tourman.2018.03.009>
- Liu, S., Law, R., Rong, J., Li, G., & Hall, J. (2013). Analyzing changes in hotel customers' expectations by trip mode. *International Journal of Hospitality Management*, 34, 359–371.
- Ma, Y., Xiang, Z., Du, Q., & Fan, W. (2018). Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep learning. *International Journal of Hospitality Management*, 71, 120–131. <https://doi.org/10.1016/j.ijhm.2017.12.008>
- Mariani, M. M., & Borghi, M. (2020). Online Review Helpfulness and Firms' Financial Performance: An Empirical Study in a Service Industry. *International Journal of Electronic Commerce*, 24(4), 421–449. <https://doi.org/10.1080/10864415.2020.1806464>
- Mariani, M. M., Ek Styven, M., & Teulon, F. (2021). Explaining the intention to use digital personal data stores: An empirical study. *Technological Forecasting and Social Change*, 166, 120657. <https://doi.org/10.1016/j.techfore.2021.120657>

- Mariani, M. M., & Nambisan, S. (2021). Innovation Analytics and Digital Innovation Experimentation: The Rise of Research-driven Online Review Platforms. *Technological Forecasting and Social Change*, 172, 121009. <https://doi.org/10.1016/j.techfore.2021.121009>
- Mariani, M., & Borghi, M. (2021b). Are environmental-related online reviews more helpful? A big data analytics approach. *International Journal of Contemporary Hospitality Management*, 33(6), 2065–2090. <https://doi.org/10.1108/IJCHM-06-2020-0548>
- Mariani, M. M., & Borghi, M. (2018). Effects of the Booking.com rating system: Bringing hotel class into the picture. *Tourism Management*, 66, 47–52. <https://doi.org/10.1016/j.tourman.2017.11.006>
- Mariani, M. M., & Borghi, M. (2021a). Environmental discourse in hotel online reviews: A big data analysis. *Journal of Sustainable Tourism*, 29(5), 829–848. <https://doi.org/10.1080/09669582.2020.1858303>
- Mariani, M. M., & Fosso Wamba, S. (2020). Exploring how consumer goods companies innovate in the digital age: The role of big data analytics companies. *Journal of Business Research*, 121, 338–352. <https://doi.org/10.1016/j.jbusres.2020.09.012>
- Mariani, M. M., Borghi, M., & Gretzel, U. (2019a). Online reviews: Differences by submission device. *Tourism Management*, 70, 295–298. <https://doi.org/10.1016/j.tourman.2018.08.022>
- Mariani, M. M., Borghi, M., & Kazakov, S. (2019b). The role of language in the online evaluation of hospitality service encounters: An empirical study. *International Journal of Hospitality Management*, 78, 50–58. <https://doi.org/10.1016/j.ijhm.2018.11.012>
- Mariani, M. M., Di Felice, M., & Mura, M. (2016). Facebook as a destination marketing tool: Evidence from Italian regional destination management organizations. *Tourism Management*, 54, 321–343. <https://doi.org/10.1016/j.tourman.2015.12.008>
- Mariani, M. M., Mura, M., & Di Felice, M. (2018). The determinants of Facebook social engagement for national tourism organizations' Facebook pages: A quantitative approach. *Journal of Destination Marketing & Management*, 8, 312–325. <https://doi.org/10.1016/j.jdmm.2017.06.003>
- Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: A systematic literature review. *International Journal of Contemporary Hospitality Management*, 30(12), 3514–3554. <https://doi.org/10.1108/IJCHM-07-2017-0461>
- Marine-Roig, E., & Clave, S. A. (2015). Tourism analytics with massive user generated content: A case study of Barcelona. *Journal of Destination Marketing & Management*, 4(3), 162–172. <https://doi.org/10.1016/j.jdmm.2015.06.004>
- Masau, P., & Prideaux, B. (2003). Sustainable tourism: A role for Kenya's hotel industry. *Current Issues in Tourism*, 6(3), 197–208. <https://doi.org/10.1080/13683500308667953>
- Mehraliyev, F., Kirilenko, A. P., & Choi, Y. (2020). From measurement scale to sentiment scale: Examining the effect of sensory experiences on online review rating behavior. *Tourism Management*, 79, 104096. <https://doi.org/10.1016/j.tourman.2020.104096>
- Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A big data analytics method for tourist behaviour analysis. *Information & Management*, 54(6), 771–785. <https://doi.org/10.1016/j.im.2016.11.011>
- Mirzaalian, F., & Halpenny, E. (2019). Social media analytics in hospitality and tourism: A systematic literature review and future trends. *Journal of Hospitality and Tourism Technology*, 10(4), 764–790.
- Moro, S., Rita, P., Ramos, P., & Esmerado, J. (2019). Analysing recent augmented and virtual reality developments in tourism. *Journal of Hospitality and Tourism Technology*, 10(4), 571–586. <https://doi.org/10.1108/JHTT-07-2018-0059>
- Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory and Practice*, 41(6), 1029–1055. <https://doi.org/10.1111/etap.12254>
- Nambisan, S., & Baron, R. A. (2007). Interactions in virtual customer environments: Implications for product support and customer relationship management. *Journal of Interactive Marketing*, 21(2), 42–62. <https://doi.org/10.1002/dir.20077>
- Nambisan, S., & Nambisan, P. (2008). How to profit from a better virtual customer environment. *MIT Sloan Management Review*, 49(3), 53.
- Oklevik, O., Gössling, S., Hall, C. M., Steen Jacobsen, J. K., Grøtte, I. P., & McCabe, S. (2019). Overtourism, optimisation, and destination performance indicators: A case study of activities in Fjord Norway. *Journal of Sustainable Tourism*, 27(12), 1–21.
- Paldino, S., Bojic, I., Sobolevsky, S., Ratti, C., & González, M. C. (2015). Urban magnetism through the lens of geo-tagged photography. *EPJ Data Science*, 4(1), 1–17. <https://doi.org/10.1140/epjds/s13688-015-0043-3>
- Pan, B., Xiang, Z., Law, R., & Fesenmaier, D. R. (2011). The dynamics of search engine marketing for tourist destinations. *Journal of Travel Research*, 50(4), 365–377.
- Pedersen, E. R., & Neergaard, P. (2006). Caveat emptor – let the buyer beware! Environmental labelling and the limitations of 'green' consumerism. *Business Strategy and the Environment*, 15(1), 15–29. <https://doi.org/10.1002/bse.434>
- Peiró-Signes, A., Segarra-Oña, M. D. V., Verma, R., Mondéjar-Jiménez, J., & Vargas-Vargas, M. (2014). The impact of environmental certification on hotel guest ratings. *Cornell Hospitality Quarterly*, 55(1), 40–51. <https://doi.org/10.1177/1938965513503488>

- Pencle, N., & Mălăescu, I. (2016). What's in the words? Development and validation of a multidimensional dictionary for CSR and application using prospectuses. *Journal of Emerging Technologies in Accounting*, 13 (2), 109–127. <https://doi.org/10.2308/jeta-51615>
- Peng, G., Liu, Y., Wang, J., & Gu, J. (2017). Analysis of the prediction capability of web search data based on the HEDC method—prediction of the volume of daily tourism visitors. *Journal of Systems Science and Systems Engineering*, 26(2), 163–182. <https://doi.org/10.1007/s11518-016-5311-7>
- Prayag, G., Chowdhury, M., Prajogo, D., Mariani, M., & Guizzardi, A. (2022). Residents' perceptions of environmental certification, environmental impacts and support for the world expo 2015: the moderating effect of place attachment. *International Journal of Contemporary Hospitality Management*, 34(3), 1204–1224. <https://doi.org/10.1108/IJCHM-06-2021-0824>
- Radojevic, T., Stanisic, N., & Stanic, N. (2015). Ensuring positive feedback: Factors that influence customer satisfaction in the contemporary hospitality industry. *Tourism Management*, 51, 13–21. <https://doi.org/10.1016/j.tourman.2015.04.002>
- Raub, S., & Martin-Rios, C. (2019). Think sustainable, act local" – a stakeholder-filter-model for translating SDGs into sustainability initiatives with local impact. *International Journal of Contemporary Hospitality Management*, 31(6), 2428–2447.
- Revinat (2018). *Global Hotel Reputation Benchmark Report 2017*. Retrieved 12 January 2022, from <https://learn.revinat.com/hospitality-research-studies/global-hotel-reputation-benchmark-report-2017>
- 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*, 15(11), 2537.
- Rosario, A. B., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, 53(3), 297–318. <https://doi.org/10.1509/jmr.14.0380>
- Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. Boston Consulting Group. Retrieved January 12, 2022, from https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx
- Salimath, M. S., & Chandna, V. (2021). Sustainable consumption and growth: Examining complementary perspectives. *Management Decision*, 59(6), 1228–1248. <https://doi.org/10.1108/MD-12-2016-0934>
- Salvatella, R. (2014). *Big Data and Tourism: New Indicators for Tourism Management*. Barcelona: Telefónica I + D and RocaSalvatella, Retrieved 12 January 2022, from [BIG-DATA-Y-TURISMO-eng-interactivo.pdf](https://www.big-data-y-turismo.com/eng-interactivo.pdf) (telefoniacatalunya.com).
- Serrano, L., Ariza-Montes, A., Nader, M., Sianes, A., & Law, R. (2021). Exploring preferences and sustainable attitudes of Airbnb green users in the review comments and ratings: A text mining approach. *Journal of Sustainable Tourism*, 29(7), 1134–1152. <https://doi.org/10.1080/09669582.2020.1838529>
- Straughan, R. D., & Roberts, J. A. (1999). Environmental segmentation alternatives: A look at green consumer behavior in the new millennium. *Journal of Consumer Marketing*, 16(6), 558–575. <https://doi.org/10.1108/07363769910297506>
- Sun, S., Wei, Y., Tsui, K.-L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, 70, 1–10. <https://doi.org/10.1016/j.tourman.2018.07.010>
- Sun, T., Youn, S., Wu, G., & Kuntaraporn, M. (2006). Online word-of-mouth (or mouse): An exploration of its antecedents and consequences. *Journal of Computer-Mediated Communication*, 11(4), 1104–1127. <https://doi.org/10.1111/j.1083-6101.2006.00310.x>
- Sutherland, I., Sim, Y., Lee, S. K., Byun, J., & Kiatkawsin, K. (2020). Topic modeling of online accommodation reviews via latent dirichlet allocation. *Sustainability*, 12(5), 1821. <https://doi.org/10.3390/su12051821>
- Thøgersen, J., Haugaard, P., & Olesen, A. (2010). Consumer responses to ecolabels. *European Journal of Marketing*, 44(11/12), 1787–1810. <https://doi.org/10.1108/03090561011079882>
- Thøgersen, J., Jørgensen, A.-K., & Sandager, S. (2012). Consumer decision making regarding a "green" everyday product. *Psychology and Marketing*, 29(4), 187–197. <https://doi.org/10.1002/mar.20514>
- Thongplew, N., Spaargaren, G., & van Koppen, C. S. A. K. (2017). Companies in search of the green consumer: Sustainable consumption and production strategies of companies and intermediary organizations in Thailand. *Njas: Wageningen Journal of Life Sciences*, 83(1), 12–21. <https://doi.org/10.1016/j.njas.2017.10.004>
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1), 24–36. <https://doi.org/10.2307/1907382>
- UNCTAD. (2019). *Digital Economy Report 2019. Value Creation and Capture: Implications for Developing Countries*. Retrieved January 12, 2022, from <https://unctad.org/webflyer/digital-economy-report-2019>
- United Nations. (2013). Global issues: Environment. <http://www.un.org/en/globalissues/environment>
- Watts, S. (2015). Corporate social responsibility reporting platforms: Enabling transparency for accountability. *Information Technology and Management*, 16(1), 19–35. <https://doi.org/10.1007/s10799-014-0192-2>

- Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97–121. <https://doi.org/10.1509/jm.15.0413>
- 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. <https://doi.org/10.1016/j.tourman.2016.10.001>
- Yang, Y., Park, S., & Hu, X. (2018). Electronic word of mouth and hotel performance: A meta-analysis. *Tourism Management*, 67, 248–260. <https://doi.org/10.1016/j.tourman.2018.01.015>
- Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, 28(1), 180–182. <https://doi.org/10.1016/j.ijhm.2008.06.011>
- Yu, Y., Li, X., & Jai, T. M. (2017). The impact of green experience on customer satisfaction: Evidence from TripAdvisor. *International Journal of Contemporary Hospitality Management*, 29(5), 1340–1361. <https://doi.org/10.1108/IJCHM-07-2015-0371>
- Zhang, X., Yang, Y., Zhang, Y., & Zhang, Z. (2020). Designing tourist experiences amidst air pollution: A spatial analytical approach using social media. *Annals of Tourism Research*, 84, 102999. <https://doi.org/10.1016/j.annals.2020.102999>
- Zhang, Z., Zhang, Z., & Yang, Y. (2016). The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. *Tourism Management*, 55, 15–24. <https://doi.org/10.1016/j.tourman.2016.01.004>
- Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, 76, 111–121. <https://doi.org/10.1016/j.ijhm.2018.03.017>