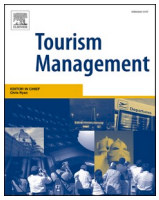
Tourism Management 100 (2024) 104821

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman



A social media analysis of travel preferences and attitudes, before and during Covid-19

Daniel Hardt, Fumiko Kano Glückstad *

Copenhagen Business School, Department of Management, Society & Communication, Dalgas Have 15, DK-2000, Frederiksberg, Denmark

ARTICLE INFO

Keywords:
Covid-19
Crisis management
Risk awareness
Travel sentiment
Travel preference
Social media
Text mining
Natural language processing

ABSTRACT

Covid-19 created tremendous uncertainty in the tourism industry; in this study, we use social media data to explore differences in the preferences and attitudes of tourism consumers, both before and during the pandemic. We use natural language processing (NLP) techniques to analyze over one million Reddit posts on travel-related subreddits. We investigate the preference for city and nature-oriented tourism in selected destinations; the analysis demonstrates that nature tourism gained interest during Covid-19 in destinations with rich nature resources, whereas city tourism lost interest in destinations known for city tourism. We also classify Reddit authors into two categories: conservation and openness, according to a psychological theory of personal values, and show that this is predictive, with openness associated with positive travel sentiment and low risk awareness. This points to the potential for value-based segmentation of travel consumers based on theoretically-grounded NLP analysis of social media data.

1. Introduction

Since the declaration of a Public Health Emergency of International Concern issued on Jan. 30, 2020 by the World Health Organization (WHO), the Covid-19 pandemic imposed severe challenges on the experience economy sector in maintaining their businesses and workplaces (Gössling, Scott, & Hall, 2020). As Covid-19 infection is closely associated with individuals' behaviors in public spaces, international tourism was severely affected by this worldwide pandemic (Chinazzi et al., 2020). The Covid-19 pandemic was a global crisis, and there were variations in domestic public health policies and international travel controls (e.g. lockdown and quarantine) implemented by authorities in different countries and regions (Collins-Kreiner & Ram, 2020; Hale et al., 2021; Mach et al., 2021; Santos, Madrid González, Haegeman, & Rainoldi, 2020); in addition, there were great variation in the reaction of individuals to the measures implemented by the country in which they reside (Glückstad, Wiil, Mansourvar, & Andersen, 2021).

One of the factors that differentiates individuals' reactions in a crisis situation is their perception of risk (Wolff, Larsen, & Øgaard, 2019), which is generally defined as an individual's assessment of the possibility of negative outcomes (Wolff et al., 2019). In the context of tourism research, Lepp and Gibson (2003) summarize four major risk factors from previous investigations: terrorism, war and political instability, health concerns, and crime. In terms of the Covid-19 and tourism, Zenker, Braun, and Gyimóthy (2021) distinguish two types of risk perceptions, one related to health (Reisinger & Mavondo, 2005; Rittichainuwat & Chakraborty, 2009) and another related to general travel activities (Floyd, Gibson, Pennington-Gray, & Thapa, 2004; Seabra et al., 2013). For preventing virus infections, people react to protect themselves and others at different levels depending on their risk perception (Parady, Taniguchi, & Takami, 2020; Shiina et al., 2020).

The risk perception of individuals is closely associated with personal factors (Lepp & Gibson, 2003, 2008) such as age, life-stage, gender, personality and value priorities in life (Glückstad et al., 2021; Kaptan, Shiloh, & Önkal, 2013) as well as travel experiences (Lepp & Gibson, 2003, 2008). Wolf, Haddock, Manstead, and Maio (2020) argue that personal values in particular can be a useful predictor, as "there is evidence linking higher conservation values and lower openness values to compliant and security-oriented behavior (Wolf et al., 2020, p. 621; see also Bardi & Schwartz, 2003; Fischer & Smith, 2006; Schwartz et al., 2017). As consumers' willingness to visit a specific destination or to experience a specific tourism product is closely connected with their perceived risk associated with the destination or product, a person who is very risk-conscious may prefer different destinations and travel products than someone who is less risk-conscious (Beirman, 2002; Lepp & Gibson, 2003, 2008; Seabra, Dolnicar, Abrantes, & Kastenholz, 2013;

* Corresponding author.
E-mail addresses: dha.msc@cbs.dk (D. Hardt), fkg.msc@cbs.dk (F.K. Glückstad).

https://doi.org/10.1016/j.tourman.2023.104821
Received 27 March 2022; Received in revised form 31 May 2023; Accepted 21 July 2023
Available online 1 August 2023

Sönmez & Graefe, 1998).

Earlier works (Dolnicar, 2005, 2006) identify market segments distinguishing the levels of perceived risk, investigating differences in the intention to travel in a situation involving risk. Whereas these earlier works use data collected through a survey, we perform similar classifications, but instead of surveys we rely on social media data. We see social media analysis as a promising alternative and complement to survey-based research. An obvious advantage of social media research is that a virtually unlimited amount of data is available because individuals autonomously create data in their natural languages. For example, we have used data from Reddit, where data is freely available, and furthermore is organized into subreddits, allowing us to target specific travel-related data. We use social media authors' posts to classify them into conservation and openness groups (Bardi & Schwartz, 2003), and we show that these classifications indeed are predictive of travel sentiment and risk awareness. These findings echo those of previous work alluded to above, but here the results are on a vastly larger scale and based on data created by authors without researchers' interference.

Another important perspective in the context of tourism research is how a destination or a tourism product is perceived by tourism consumers before and during a crisis. In the context of Covid-19, researchers have observed shifts in consumers' preferences in tourism experiences (i.e., products) during the Covid-19 pandemic (Huang, Shao, Zeng, Liu, & Li, 2021; Im, Kim, & Choeh, 2021; Marques, Guedes, & Bento, 2021). Countries severely hit by Covid-19 were associated with a negative destination image, especially by those who were concerned about health risks (Rasoolimanesh, Seyfi, Rastegar, & Hall, 2021). Hence, it is meaningful to investigate consumers' associations with destinations and tourism products, both before and during the Covid-19 pandemic. In this paper we examine preferences for city and nature tourism expressed in these social media posts, and we find that they differ in interesting ways in response to Covid-19.

Overall, the current study consists of two types of analyses, based on social media data:

1. **Preferences:** we explore consumers' preference for city vs. nature tourism products both before and during Covid-19, in different destinations, and
2. **Values and Attitudes:** We classify users in terms of two personal values: conservation and openness, based on the text of their posts in a period before Covid-19. Then we explore the extent to which these two groups differ in their travel attitudes, both before and during Covid-19.

We show how techniques in Natural Language Processing (NLP) make it possible to robustly analyze user characteristics and predict their attitudes about travel during Covid-19.

2. Literature review and contributions of this study

2.1. The impact of tourism crises on tourism products and destinations

Sönmez, Apostolopoulos, and Tarlow (1999, pp. 14-15) define a tourism crisis as an occurrence which can "damage a tourist destination's overall reputation for safety, attractiveness, and comfort by negatively affecting visitors' perceptions of that destination." Whereas a risk perceived by tourism consumers negatively impacts the attractiveness of a destination or a tourism product (Hu & Ritchie, 1993), a risk assessed by tourism managers typically influences their crisis planning and response strategies (Wang & Ritchie, 2012). The impact of a crisis depends on its nature, intensity, and scale (Backer & Ritchie, 2017). As seen during the Severe Acute Respiratory Syndrome (SARS) outbreak in Asia in 2003, travel activities - in particular international tourism - are inherently connected to the global spread of an infectious disease (Henderson & Ng, 2004; Mason, Grabowski, & Du, 2005; McKercher & Chon, 2004; Novelli, Gussing Burgess, Jones, & Ritchie, 2018; Washer, 2004). Several studies investigate the impacts of infectious diseases (e.g. SARS and Covid-19) on consumers' risk-averse actions (Im et al., 2021), travel intentions (Floyd et al., 2004; Kozak, Crotts, & Law, 2007; Sánchez-Cañizares, Cabeza-Ramírez, Muñoz-Fernández, & Fuentes--García, 2021; Zenker et al., 2021), travel avoidance (Cahyanto, Wiblishauser, Pennington-Gray, & Schroeder, 2016), tourism preferences (Huang et al., 2021; Marques et al., 2021), and travel behaviors (Li, Nguyen, & Coca-Stefaniak, 2021; Neuburger & Egger, 2021).

Consumers' risk perception of a specific crisis is one factor that affects their attitudes and preferences concerning potential destinations and tourism products (Cahyanto et al., 2016; Floyd et al., 2004; Lepp & Gibson, 2003; Neuburger & Egger, 2021; Reichel, Fuchs, & Uriely, 2007; Sánchez-Cañizares et al., 2021; Taylor & Toohey, 2007; Zenker et al., 2021; Zheng, Luo, & Ritchie, 2021). Novelli et al. (2018) emphasize that the outbreak of an infectious disease creates a spillover effect to neighboring destinations that have not been directly affected by the outbreak (Cavlek, 2002; Henderson, 2007; Ritchie, Crotts, Zehrer, & Volsky, 2014). Such a spillover effect has also been known as a generalization effect, i.e., consumers tend to associate surrounding destinations of a specific crisis and generalize them as unsafe (Enders, Sandler, & Parise, 1992; Lepp & Gibson, 2008).

There are numerous works arguing that tourism consumers, being generally concerned about public hygiene and social distancing, tend to avoid a destination or a tourism product that is directly or indirectly associated with health-risks. For example, during Covid-19, travelers tend to prefer low density rural destinations (Marques et al., 2021) with rich open areas enabling 'untact (no contact)' activities (Lee & Lee, 2020). A series of studies conducted in February 2020 by Huang et al. (2021) also report that Covid-19 reduced Chinese nationals' willingness to travel to geographically and culturally distant destinations and made nature-based rural destinations a more favorable place to visit. Similarly, Marques et al. (2021) demonstrate that consumers' googling about rural domestic destinations in Portugal substantially increased during the summer 2020. Similar reports are also identified in Indonesia (Wachyuni & Kusumaningrum, 2020) and the Czech Republic (Vaishar & Šťastná, 2022). These studies document that tourism consumers' travel preferences have shifted during Covid-19 to rural destinations with rich natural resources, which potentially minimize consumers' perceived health-risks.

2.2. Media coverage, risk perception and negative destination image formation

Consumers' perceived risk associated with a destination is also affected by media imagery of a destination (Cahyanto et al., 2016; Mason et al., 2005; Rasoolimanesh et al., 2021; Yang, Isa, & Ramayah, 2021). In particular, negative media coverage of a crisis in a specific destination affects individuals' perceptions of that destination (Kapuściński & Richards, 2016; Novelli et al., 2018; Schroeder & Pennington-Gray, 2014) and the formation of destination images (Baloglu & McCleary, 1999; Beerli & Martín, 2004; Kozak et al., 2007; Mazanec & Strasser, 2007). For example, when media identifies a certain destination - such as China or Italy - as an epicenter of the Covid-19 pandemic, tourism consumers most likely perceive these destinations as 'unsafe', thereby avoid these destinations (Huang et al., 2021). Similarly, Glückstad, Schmidt, and Mørup (2020) demonstrate that a group of tourism consumers who associated uncertainty with a politically unstable destination in the Middle East also associated Paris with risk-related attributes just after the international media coverage of the terror attack in November 2015, and argue that a destination image held by tourism consumers is dynamically revised when they become exposed to negative media coverage about a crisis (Glückstad et al., 2020). Cherifi, Smith, Maitland, and Stevenson (2014) call this a naïve image of a destination (Luo & Zhai, 2017; Selby, 2004), typically created without a direct experience, but shaped via a stereotypical image

(Jenkins, 1999; Mason et al., 2005) reported by media. In contemporary society, social media also plays an important role in shaping consumers' naïve images about a destination or a product (Luo & Zhai, 2017; Mizrachi & Fuchs, 2016; Möller, Wang, & Nguyen, n. d.; Schroeder & Pennington-Gray, 2014; Vu, Li, & Law, 2020). However, Abubakar, Ilkan, Meshall Al-Tal, and Eluwole (2017); Han et al. (2021); Rasoolimanesh et al. (2021) argue that when tourists associate a destination with trustworthy social capital, the trust serves to minimize their uncertainty and perceived risk. Finally, Kapuściński and Richards (2016) apply a framing theory of media effects to explain how psychographic characteristics moderate the effect of media coverage on their destination risk perception. Moreover, the framing theory argues that quality of media coverage also affects risk perception of information receivers (Kapuściński & Richards, 2016). From this point, the content analysis of news media by Mach et al. (2021) demonstrated substantial differences in reporting the event of Covid-19 across newspapers in USA, UK and Canada.

These studies indicate the importance and the complexity of health-related crisis communication targeting various types of tourism consumers who might perceive risks associated with a specific destination or a tourism product at different levels (Floyd et al., 2004) during a pandemic scenario, such as Covid-19 (Lawton & Page, 1997; Villacé-Molinero, Fernández-Muñoz, Orea-Giner, & Fuentes-Moraleda, 2021).

2.3. Personal values and risk perception

For the purpose of tourism crisis management, it is important to identify consumer segments that possess crisis-resistant characteristics. One way to identify an attractive segment is to classify tourism consumers according to the level of their perceived risk (Dolnicar, 2005, 2006). For example, Hajibaba, Gretzel, Leisch, and Dolnicar (2015) identified respondents who experienced a crisis in the past, and classified them into six clusters based on their actual reaction (i.e., cancel or continue the trip) to the crisis event. Their study emphasizes the importance of identifying and understanding a crisis-resistant segment for the purpose of marketing and communication. Liu, Schroeder, Pennington-Gray, and Farajat (2016) classifies US tourists into four clusters based on two dimensions: avoidance (high perceived risk - low efficacy beliefs), responsive (high risk - high efficacy), indifference (low risk - low efficacy), and proactive (low risk – high efficacy). Their findings demonstrate that the proactive segment maintained positive association with and higher intention to visit Jordan. Handler (2016) segments Taiwanese travelers according to various types of risk perception of Japan 2.5 years after the Fukushima disaster in 2011. Their study identifies five distinctive clusters (Apprehensive travelers, Informed travelers, Health-conscious travelers, Carefree travelers, and Fearful travelers). Interestingly, these clusters involve distinct type of concerns, while most of them maintained a willingness to visit Japan. Furthermore, Kozak et al. (2007) classifies individuals residing in three types of countries based on Hofstede's uncertainty avoidance index (Hofstede, 1980). Their findings show that international travelers are generally sensitive to risks involved in a variety of contexts such as infectious disease, terrorist attacks, and natural disasters, and culture-specific segments differentiate the level of risk perception.

Consumers' travel and health-related risk perception are particularly influenced by personal factors (Floyd et al., 2004; Glückstad et al., 2020; Kozak et al., 2007; Parady et al., 2020; Reichel et al., 2007; Reisinger & Mavondo, 2005; Rittichainuwat & Chakraborty, 2009; Schroeder & Pennington-Gray, 2014; Seabra et al., 2013; Shiina et al., 2020; Taylor & Toohey, 2007; Wolff et al., 2019). While demographic variables such as nationality, age, gender, and life stages may be useful indicators to predict different levels of travel risk perception held by consumers, earlier studies (Floyd et al., 2004; Lepp & Gibson, 2003, 2008; Rasoolimanesh et al., 2021; Rittichainuwat & Chakraborty, 2009; Sönmez & Graefe, 1998) indicate that previous travel experience can be a moderator of consumers' risk perception and become a significant predictor of intention to travel. In the context of health-related risks, in particular, in the context of epidemic and pandemic, personality and value priorities in life can be another useful predictor (Bardi & Schwartz, 2003; Fischer & Smith, 2006; Schwartz et al., 2017; Wolf et al., 2020). Specifically, Wolf et al. (2020) argue for the usefulness of the Schwartz's theory of ten basic human values (Schwartz, 1992, 2006; Schwartz et al., 2012). In Schwartz's theory (1992; 2006) ten basic values (self-direction, stimulation, hedonism, achievement, power, tradition, conformity, security, benevolence, and universalism) are hierarchically structured in a quasi-circumplex model (Fig. 1). These ten values are placed in a two dimensional space where the vertical dimension refers to Anxiety-Free/Growth versus Anxiety-Avoidance/Self-Protection values whereas the horizontal dimension refers to Personal-focus versus Social-focus values. The model defines four higher-order values: openness to change, subsuming self-direction, stimulation and hedonism; conservation, consisting of conformity, tradition and security; self-transcendence, with universalism and benevolence; and self-enhancement, which involves power and achievement. In other words, the value theory by Schwartz highlights that the openness to change oriented people possess risk-free and personal-focus orientation characterized as an internal locus of control (self-direction), seeking for thrills (stimulation) and pleasure (hedonism) which is consistent with the construct of sensation seeking (Zuckerman & Aluja, 2015). On the other hand, the conservation oriented people can be characterized as social-focused and risk-aversion.

In the current study we focus on the higher-order values of openness and conservation, as we see a natural connection with the issues of risk sensitivity and travel interest. The other two higher-order values, self-enhancement and self-transcendence, are perhaps less clearly connected to these issues, although that is ultimately a question for empirical investigation.

In recent work Kaptan et al. (2013) examine the relationship between personal values and risk perceptions regarding a crisis situation such as a terror attack. They examine risk perception from both an emotional perspective (e.g., fear, helplessness, insecurity, anxiety) and a cognitive perspective (perceived likelihood and severity) in Turkey and Israel, and they show that, while the conservation value is positively correlated with risk from emotional perspectives, the openness to change value is negatively correlated with risk from both emotional and cognitive perspectives in both countries. However, the correlations between the two perspectives of risk perception and the self-transcendence/self-enhancement values are only observed in either of the countries in their studies. Further research is needed to explore the relevance of values to a variety of travel-related attitudes and behaviors.

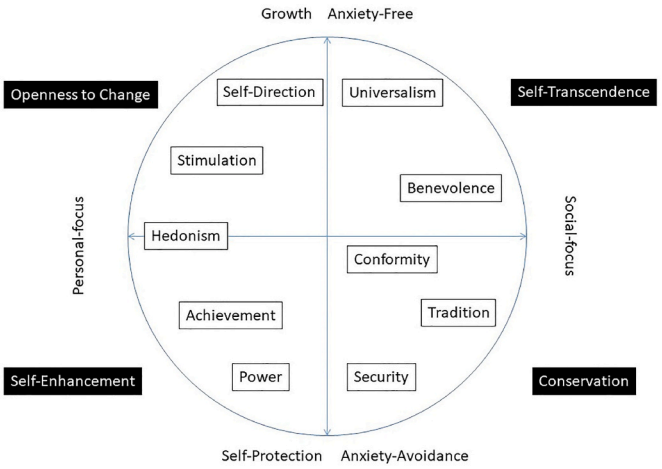


Fig. 1. Schwartz's theory of ten basic human values (based on Schwartz et al., 2012).

2.4. Social media analysis and travel attitudes

Most investigations of tourists' risk perception, destination image, and travel intention have been based on data collected via surveys (Floyd et al., 2004; Hajibaba et al., 2015; Seabra et al., 2013; Sönmez & Graefe, 1998; Zenker et al., 2021). Some recent works have implemented social media analysis using various text mining techniques. For example, Govers, Go, and Kumar (2007) use neural networks to conduct content analysis of qualitative data collected via an online survey consisting of open-ended questions. Their study demonstrated the usability of neural network techniques to extract respondents' perceived images from qualitative text data in several case studies. Li, Lin, Tsai, and Wang (2015) use text mining and semantic networks to analyze a blog sample collected from the Chinese social media platform ctrip.com, and investigate the relationships between blog users' cognitive/affective images and their image formation about a destination. Arefieva, Egger, and Yu (2021) combine three disciplines – semiotics, marketing, and data sciences – in their study to classify textual information based on Instagram photographs, and extract several destination image clusters. Their study tests various machine learning techniques and recommends that k-means clustering of document-term matrices is the best approach. Bjørkelund, Burnett, and Nørvåg (2012) present an opinion mining approach combining the result of sentiment analysis of textual contents and Google Maps to visualize opinions based on hotel review sites.

Some studies focus on the relationship between experiences of hotel customers and their satisfaction. For example, Guo, Barnes, and Jia (2017) demonstrate that linguistic analysis of 265,544 online reviews of 25,670 hotels located in 16 countries using Latent Dirichlet Analysis (LDA) revealed 19 key dimensions that are useful for hotels to improve their customer relations. Xiang et al. (2015a, 2015b) investigate the experiences of hotel guests and its association with their satisfaction, by applying a text analytic approach to 60,000 hotel reviews collected from Experiea.com. Their findings indicate a strong association between guest experience and satisfaction.

Platforms used in social media analysis also vary in these studies. For example, Shimada, Inoue, Maeda, and Endo (2011) conduct a positive-negative classification task of 116 tweets related to a destination or a tourism event. Claster, Dinh, and Cooper (2010) use 70 million tweets to analyze sentiment about Cancun, Mexico, while Claster, Dinh, and Cooper (2010) collect 80 million tweets to look at sentiment about Thailand. Finally, Chen, Hsieh, Mahmud, and Nichols (2014) combine an online questionnaire sent to Reddit users with their Reddit posts and analyze words used by 799 Reddit users who prioritize different personal values.

In the context of tourism crisis management, Gkritzali (2017) conduct a longitudinal study of the sentiment of a crisis-affected destination, Athens, hit by the financial crisis. They perform sentiment analysis of online conversations found on TripAdvisor's Athens Travel Forum, and argue that the evolution of online sentiment about Athens observed through the Forum can be an important indicator of destination image used by researchers and DMOs during the recovery process of a crisis. Another study by Mizrachi and Fuchs (2016) involves a thematic analysis of 200 posts in TripAdvisor forums about traveling to Ebola-free African countries. Their analysis extracts three thematic topics: "Positive Thinking and Encouragement; Knowledge Development and Preparation; and Personal Risk Assessment" (Mizrachi & Fuchs, 2016, p. 59). Content analysis of social media posts is also used for investigating consumers' attitudes and risk awareness during the Covid-19 pandemic. In Pantano, Priporas, Devereux, and Pizzi (2021), 15,000 tweets collected from UK, Italy, and Spain in April 2020 were analyzed by use of text classification techniques, i.e., the content extraction and phrases extraction that computes word co-occurrence of keywords and similarity measures of tweets. Their systematic content analysis of the tweets shows that consumers indicate a sense of escapism during Covid-19 and show a willingness to spend in enjoying freedom in the post-pandemic period. These studies suggest that social media content analysis and sentiment analysis can provide insight into travel-related consumer attitudes and concerns during a crisis.

2.5. Contributions of this study

The Covid-19 pandemic, with its disruption of tourism, underscores the need to assess the sentiments, preferences, and risk awareness of tourism consumers. Recent work suggests that the preference of tourism consumers shifted during Covid-19 to nature tourism that enables them to minimize uncertainty in the health-related risks (Huang et al., 2021; Marques et al., 2021); the media imageries affect perceptions of a destination at different levels (Kapuściński & Richards, 2016; Rasoolimanesh et al., 2021; Yang, Isa, & Ramayah, 2021); tourism consumers vary systematically along these dimensions, in a way that depends on individuals' personal values defined by Schwartz (2007) (Kaptan et al., 2013). The importance of investigating relations between personal values and the sentiments, preference and risk awareness of tourism consumers has received some limited empirical support, primarily based on user surveys (Huang et al., 2021; Kaptan et al., 2013; Lee & Lee, 2020; Marques et al., 2021; Vaishar & Šťastná, 2022; Wachyuni & Kusumaningrum, 2020). In the current study, we propose an alternative empirical basis for assessing consumers – instead of posing explicit questions to a small selection of individuals, we use social media data which contains a virtually limitless source of travel-relevant text data freely contributed by individuals. Although there are important differences in the manner and extent to which different subgroups use social media, we suggest that social media analysis represents a valuable, promising alternative to the survey-based approach.

While there are some studies that use social media data to analyze consumer attitudes (Gkritzali, 2017; Mizrachi & Fuchs, 2016; Pantano et al., 2021), up to this point they have been limited in scope and techniques. The current study is based on over one million posts collected from a diverse collection of travel-related subreddits. Furthermore, this study applies sophisticated techniques for semantic analysis, based on word embeddings, making it possible to use individuals' posts to quantitatively assess personal values as well as attitudes and preferences. Furthermore, word clouds are used to show how this vast collection of text data provides rich qualitative insight into mental images evoked by different destinations and travel products.

Our study poses two research questions (RQs) that will be addressed separately as study 1 and study 2, as follows:

- **RQ1 (preferences): How do preferences for city or nature tourism change before and during Covid-19 for selected destinations?**
- **RQ2 (values and attitudes): What type of tourists show positive sentiment and low risk awareness before and during Covid-19?**

To address RQ2, we classify users according to the personal values of conservation and openness, and explore the extent to which these classifications reflect differing travel attitudes, both before and during Covid-19.

We first define destinations and tourism products to be investigated in our study. As pointed out by Rasoolimanesh et al. (2021), destinations vary in terms of their perceived trustworthiness, which in turn relates to consumers' uncertainty and perceived risk. To define a baseline, we select destinations listed as trustworthy countries - USA, UK, Canada, Australia, France, Germany, Italy, Spain, and Japan – based on Culligan, Dubber, and Lotten (2014). Our study also investigates two different tourism products: nature tourism and city tourism, inspired by Huang et al. (2021); Lee and Lee (2020); Marques et al. (2021). Addressing these destinations and tourism products, we define the following propositions listed below:

1. **Tourism products before and during Covid-19:** During Covid-19, the relative preference of nature tourism versus city tourism

increased compared to the relative preference of nature versus city tourism before Covid-19.

2a. **Sentiment before and during Covid-19:** Consumers' sentiment about traveling during Covid-19 decreased from before Covid-19, both for conservation and openness consumers.

2b. **Sentiment expressed by two types of consumers:** Both before and during Covid-19, consumers classified with the openness value express more positive sentiment about traveling than those classified with the conservation value.

3a. **Risk awareness before and during Covid-19:** Consumers' risk awareness during Covid-19 increased from before Covid-19, both for conservation and openness consumers.

3b. **Risk awareness expressed by two types of consumers:** Both before and during Covid-19, consumers classified with the openness value express less risk awareness than those classified with the conservation value.

3. Methods

Fig. 2 gives an overview of the methods used to address the two research questions.

3.1. Data collection

We collect all Reddit posts from 14 travel-oriented subreddits, from the beginning of 2016 until January 16, 2021. The number of posts for each of the travel-oriented subreddits is shown in Table 1. There was a total of 1,183,504 posts by 435,344 authors. We create two subsets of posts, for before Covid-19 and during Covid-19. We define the *before Covid-19* category as posts made between Jan. 1, 2016–March 10, 2020, and the *during Covid-19* category consists of posts from March 11, 2020 until Jan. 1, 2021. March 11, 2020 was selected as the dividing point, as this is the date on which the WHO declared Covid-19 a pandemic.[1] As shown in Table 2, there are 1,024,181 posts by 420,298 authors posted before the Covid-19 pandemic and 159,322 posts by 15,046 authors posted during the Covid-19 pandemic.[2]

For study 1, we analyze the posts before and during Covid-19 in terms of preferences expressed for city and nature tourism products as shown in Fig. 2.

3.2. Classification of authors

For study 2, we wish to classify authors in terms of travel-related values, based on analysis of the text content in their posts. We use a psychographic classification approach based on the Schwartz's theory of basic human values (Schwartz, 2006, 2007; Schwartz et al., 2012), and classify authors into the openness and the conservation groups. Previously, such classifications have been done using surveys filled out by participants. Here we identify authors and analyze their posts to perform this classification. To see if such psychographic classifications can be reliably associated with travel attitudes, we divide the posts before Covid-19 into two equal sized groups as shown in Table 3. The data is structured as follows:

- **Before1:** all posts between Jan. 1, 2016 and May 24, 2018
- **Before2:** all posts between May 25, 2018 and March 10, 2020
- **During:** all posts between March 11, 2020 and Jan. 1, 2021

Unlike in study 1, we divide the before period into before1 and before2. This makes it possible to classify authors based on posts made in before1 and then compare their posts in before2 and during. This allows us to examine the predictiveness and stability of the classifications made on the basis of text posts, and we do this in two subsequent periods: both a "normal" period, and a subsequent period of crisis.

After creating these three sets of posts, we group posts by author in before1. Based on the descriptions of these psychographic types in (Schwartz et al., 2012) we create the following word lists for the two categories:

- **Openness:** stimulation, excitement, novelty, challenge, daring, adventurous, thrill, hedonism, pleasure, sensual, enjoy, fun, luxurious, indulgent, freedom, creative, curious, independent, extrovert
- **Conservation:** conservative, conformity, obedient, discipline, polite, compliant, restraint, security, safe, safety, orderly, clean, protect, careful, traditional, orthodox, stable, conventional, moderate, interdependent, introvert

We determine word embedding vectors for each keyword on the list, using the word2vec package (Mikolov, Chen, Corrado, & Dean, 2013), with word embeddings produced based on the GoogleNews corpus (Google Code Archive, 2013). Word embeddings are widely used to capture word relations such as semantic similarity. The word2vec package produces embeddings that are dense real-valued vectors, with a dimension of 300.[3] A word embedding is a representation of an individual word; as Almeida and Xexéo (2019, p. 1) point out, "word embeddings … in addition to encoding surprisingly good syntactic and semantic information, have been proven useful as extra features in many downstream NLP tasks." Word embeddings can also be used to represent collections of words: this can be done using various operations from vector arithmetic. The most typical operation is to construct an average embedding (Elsaadawy, Torki, & Ei-Makky, 2018; Sharma & Daniels, 2020). We use this operation to construct an average for the two categories of conservation and openness, based on the word lists given above.

Next, we similarly create an average embedding for each author, based on the collection of posts made by that author. Recall that we have a before1, before2 and during Covid-19 collection for each author. We perform our psychographic classification based on the before1 text collection. For each author, we create an average embedding of their before1 text collection. Then we compute the similarity of this embedding to the embedding for conservation, and we also compute its similarity to the embedding for openness. We use the metric of cosine similarity, which is the most widely used method for determining semantic similarity of word embeddings (Levy, Goldberg, & Ido, 2015, Orkphol & Yang 2019). Cosine similarity is a general method for determining the similarity of two vectors, based on a computation of the dot product of the vectors. Values range from −1 to 1.0, with 1.0 indicating that two vectors are identical; lower values reflect lower degrees of similarity. We now have, for each author, a measure of the author's similarity to openness and conservation values. Finally, we subtract the openness similarity from the conservation similarity. The resulting values, which range from −1.0 to +1.0, are such that higher values represent greater similarity to conservation and less similarity to openness; lower values represent greater similarity to openness and less similarity to conservation.[4]

As seen in Table 3, the classification of authors is based on 214,003 authors in the before1 period. We then restrict attention to authors who posted in all three periods, giving a total of 3,093 authors. Table 4

[1] ℙCOVID-19 pandemic", Wikipedia.

[2] Note that the EarthPorn subreddit has nothing to do with pornography, but rather, as described on The Definitive List of Travel Subreddits, it contains "absolutely breathtaking landscape photography showcased by talented photographers."

[3] Note that the number of authors is the same for all three periods, because we restrict our attention to the 3,093 who posted during all three periods.

[4] Note that our data collection ends on January 1, 2021. In future work, we would like to collect data on the later portions of the pandemic, as well as data as the world begins to emerge.

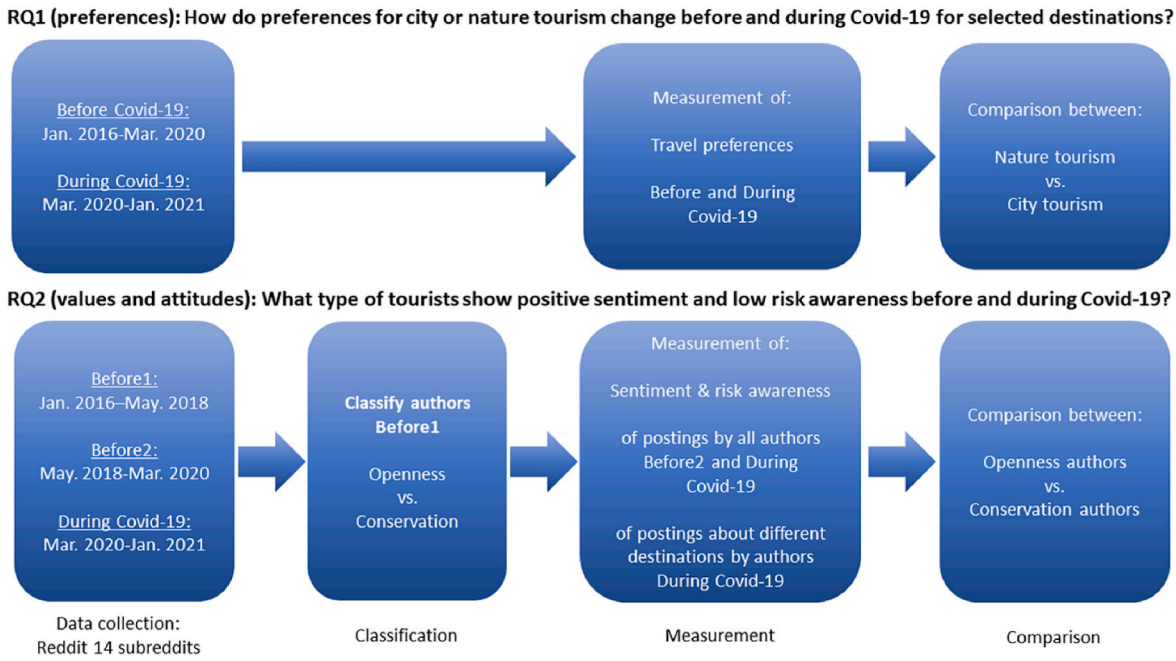


Fig. 2. Overview of the studies (RQ1: Study1, RQ2: Study2).

Table 1
Subreddits and number of posts.

Subreddit	Number of posts
Travel	486,097
EarthPorn	464,186
SoloTravel	63,931
CampingandHiking	59,379
Backpacking	48,128
Cruise	18,646
Flights	11,510
TravelHacks	8,589
Adventures	8,112
Travelphotos	6,314
TravelBlog	3,472
Wanderlust	2,720
TravelNoPics	1,340
RemotePlaces	1,080
Total	**1,183,504**

Table 2
Number of posts and authors before and during Covid-19.

	Number of posts	Number of authors
Before Jan. 1, 2016–Mar. 10, 2020	1,024,181	420,298
During Mar. 11, 2020–Jan. 1, 2021	159,321	78,704

presents summary statistics of the conservation-openness similarity scores for the posts in the before1 period posted by the 3,093 authors. The minimum score (−0.29) is the value most similar to openness and the maximum score (0.22) is the value most similar to conservation. The mean and median values are quite similar, and the skewness score indicates that the distribution of values is quite symmetrical. Accordingly, we divide these values into two halves of equal size: authors in the top half we identify as conservation authors, and authors in the lower half are identified as openness authors.

3.3. Measurement - before and during Covid-19

Following the procedure in Fig. 2, we measure travel interests in the two types of tourism products, nature tourism and city tourism, as well as sentiment and risk awareness of the two types of authors (openness and conservation) both before and during Covid-19. We further calculate destination-specific travel preferences before and during Covid-19 by selecting authors who mention the specific destinations listed below:

Table 3
Number of posts and authors for study 2.

	Number of posts	Number of authors
All authors		
Before 1 Jan. 1, 2016 and May 24, 2018	467,534	214,003
Before 2 May 25, 2018–Mar. 10, 2020	467,544	230,218
During Mar. 11, 2020–Jan. 1, 2021	159,321	78,704
Authors common across the three periods		
Before 1 Jan. 1, 2016 and May 24, 2018	25,198	3,093
Before 2 May 25, 2018–Mar. 10, 2020	30,190	3,093
During Mar. 11, 2020–Jan. 1, 2021	11,117	3,093

- **The selected destinations:** USA, UK, Canada, Australia, France, Germany, Italy, Spain, Japan

Study 1: Preference of Tourism Products.

We first measure degrees of interest in the two travel products using all posts. We consider *city* and *nature* as two particularly relevant tourism products, especially during a health crisis. We measure interest in city and nature using a technique of average embeddings based on word lists. Below are the lists of unique words representing city and nature inspired by Baloglu and McCleary (1999); Beerli and Martín (2004):

- **City:** city, urban, crowd, museum, shopping, dancing, restaurant, bar, casino, entertainment, discotheques, lively
- **Nature:** nature, rural, sunset, beach, mountain, lake, tree, countryside, trekking, cleanliness, quietness

The scores for the two travel products are computed for all posts as well as for posts mentioning the specific destinations. Our approach here

Table 4
Summary statistics of the openness-conservation similarity scores.

Means, median and ranges of Cosine similarity scores Before 1										
N	Mean	Median	Variance	Std. Deviation	Std. Error of Mean	Minimum	Maximum	Range	Kurtosis	Skewness
3093	−.0454	−.0428	.003	.05104	.00092	−.29	.22	.51	1.189	−.182

is similar to that used for classifying texts as similar to average embeddings for the concepts of conservation and openness. We determine average embeddings for city and nature based on the above word lists. Then we use cosine similarity to determine the similarity of a text to the city embedding and to the nature embedding, and we consider this similarity to be a measure of the interest expressed in the text, with values ranging from −1.0 (minimal interest) to +1.0 (maximum interest).

Study 2: Sentiment and Risk Awareness.

To measure sentiment, we use the Sentiment Intensity Analyzer from the Natural Language Toolkit (Team, 2021). For each author's posts, we compute a *compound* score, which ranges from −1.0 (most negative) to 1.0 (most positive).

To measure risk awareness of an author, we again use the average embeddings for the author's posts, as described above. We also determine the embedding for the concept of "risk" and that of "safe" by producing the embedding of each word. Then we take the difference between the similarity to the "risk" embedding and the similarity to the "safe" embedding. For ease of comparison, we normalize these results so that they range between 0 and 1.0.

3.4. Comparison

We verify statistical significance using different methods for the two tourism products and for the two author's psychographic types.

3.4.1. Study 1: Preference of Tourism Products

We first compute differences in the mean scores of interest between city and nature tourism for all posts as well as posts mentioning specific destinations both before and during Covid-19. As the means are computed based on similarity of a text to city and nature word embeddings, we conduct a paired sample test available in IBM Statistical Package for the Social Sciences version 28 (SPSS 28) to verify statistical significance and size effect (Cohen's D score) in mean differences.

3.4.2. Study 2: sentiment and risk awareness

Differences in the mean scores for sentiment and risk awareness are computed for all authors belonging to the conservation and the openness types. We also compute differences in the mean scores between the conservation and the openness authors for the selected destinations before and during Covid-19. To compare the statistical significance and size effect (Cohen's D score) in mean differences between the openness and the conservation users, we conduct an independent-samples *t*-test using SPSS 28.

To test the statistical significance and size effect in changes of the means scores before Covid-19 and during Covid-19, we conduct a paired-sample test for the two samples representing the conservation and the openness types.

3.5. Visualization

After the comparison, we select several scenarios identified in the previous analyses to visualize as word clouds.

4. Results

We now evaluate our results with the respect to the two studies we defined, concerning RQ1 addressing product preferences and RQ2 addressing personal values and differences in sentiment and risk awareness.

4.1. Study 1: preference of tourism products before and during Covid-19

The interest in city tourism and nature tourism both before and during Covid-19 is respectively shown in Table 5 and Table 6. Considering all posts (the bottom row of Table 5), the mean score for the interest in nature (0.26) is 0.4 higher than that for the interest in city (0.22) in the before Covid-19 period. This difference is even larger in the during Covid-19 period as shown in Table 6. Specifically, the mean score for the interest in nature (0.27) is 0.5 higher than that for the interest in city (0.22). The Cohen's D scores for all posts clearly indicate that the effect size in the mean differences is larger during Covid-19 (−0.56) than before Covid-19 (−0.39) (see Table 5 and Table 6).

We also present scores for posts mentioning specific destinations. An increase in preference for nature over city can be observed for nearly all destinations. As shown in Fig. 3, interest in city tourism during Covid-19 is lower than before Covid-19 in all selected destinations. On the other hand, interest in nature tourism during Covid-19 is higher than before Covid-19 in all selected destinations except USA. In particular, interest in nature tourism during Covid-19 became substantially higher in Australia, UK and Germany. Similarly, interest in city tourism during Covid-19 became substantially lower in USA, Germany and France. Cohen's D scores in Tables 5 and 6 also confirm that the size effects in the mean differences between nature and city tourism are larger during Covid-19 than before Covid-19 in all selected destinations.

Word clouds: City vs. Nature Tourism.

To gain insight into the increased preference for nature tourism, we examine word clouds for three selected destinations: Australia, Germany and France, before and during Covid-19 (Fig. 4).

For Australia, some nature related term such as "rock" and the city Melbourne are emphasized before Covid-19. During Covid-19, the term "local" becomes very prominent and other nature-oriented terms like "beach" and "coastline" emerge although the city Sydney also appeared. Similarly, for Germany and France, the capitals Berlin and Paris are emphasized before Covid-19, while during Covid-19 we see nature-oriented terms such as "forest", "trees", "water". In France the region Auvergne, known for hiking and vast forests, becomes important during Covid-19. It is important to acknowledge that we do observe some nature-oriented terms in the before Covid-19 word clouds, and some city-oriented terms in the during Covid-19 word clouds. However, what emerges is a clear change in emphasis, as the interest in nature tourism increased relative to the interest in city tourism during Covid-19.

Discussion: City vs. Nature Tourism.

Our first proposition is strongly supported: both before and during Covid-19, there is a statistically significant difference in the preference of nature tourism versus city tourism for all selected destinations except Japan before Covid-19. These relative preferences of nature tourism versus city tourism increased in all selected destinations during Covid-19 as confirmed by the Cohen's D scores in Tables 5 and 6 These findings are consistent with previous work that exhibited increased interest in traveling rural area and no-contact tourism during Covid 19 (Huang et al., 2021; Lee & Lee, 2020; Marques et al., 2021; Vaishar & Šťastná, 2022; Wachyuni & Kusumaningrum, 2020).

The countries selected in our study are destinations generally considered as trustworthy countries (Culligan et al., 2014); tourism consumers would normally associate these countries with trustworthy

Table 5
Degree of interest in city and nature tourism measured for all posts mentioning the selected destinations before Covid-19.

Before Covid-19	N	City		Nature		Mean Difference		
		Mean	SD	Mean	SD	Difference	P	Cohen's D
USA	20572	0.222	0.096	0.250	0.102	−0.029	.00***	−0.33
Canada	17872	0.193	0.095	0.255	0.116	−0.062	.00***	−0.68
Australia	12623	0.204	0.095	0.252	0.110	−0.048	.00***	−0.56
UK	14362	0.251	0.093	0.253	0.108	−0.002	.05*	−0.02
Germany	8267	0.228	0.101	0.242	0.125	−0.014	.00***	−0.17
France	11572	0.248	0.096	0.246	0.104	0.003	.00***	0.03
Italy	15220	0.240	0.093	0.262	0.110	−0.022	.00***	−0.28
Spain	7648	0.234	0.098	0.249	0.108	−0.015	.00***	−0.20
Japan	9303	0.237	0.100	0.237	0.107	0.000	.986	0.00
All	1024181	0.224	0.096	0.262	0.116	−0.037	.00***	−0.39

Table 6
Degree of interest in city and nature tourism measured for all posts mentioning the selected destinations *during Covid-19*.

During Covid-19	N	City		Nature		Mean Difference		
		Mean	SD	Mean	SD	Difference	P	Cohen's D
USA	4155	0.201	0.099	0.241	0.104	−0.040	.00***	−0.46
Canada	3705	0.190	0.095	0.264	0.121	−0.074	.03**	−0.88
Australia	1850	0.202	0.101	0.270	0.119	−0.068	.00***	−0.82
UK	2174	0.240	0.091	0.276	0.114	−0.036	.00***	−0.39
Germany	1799	0.208	0.105	0.258	0.144	−0.050	.00***	−0.57
France	1418	0.227	0.105	0.249	0.124	−0.022	.00***	−0.26
Italy	2062	0.234	0.099	0.274	0.124	−0.039	.00***	−0.48
Spain	954	0.228	0.102	0.253	0.121	−0.025	.00***	−0.33
Japan	1238	0.235	0.106	0.245	0.114	−0.010	.00***	−0.12
All	159321	0.217	0.098	0.270	0.123	−0.053	.00***	−0.56

Note1: p-values indicate statistical differences between city and nature tourism ($p < 0.01$***, $p < 0.05$**, and $p < 0.1$*).
Note2: Cohen's D scores indicate effect size (0.2 refers to small effect size, 0.5 medium effect size, and 0.8 large effect size).
Note 3: Detailed statistics are found in Appendices 1 and 2.

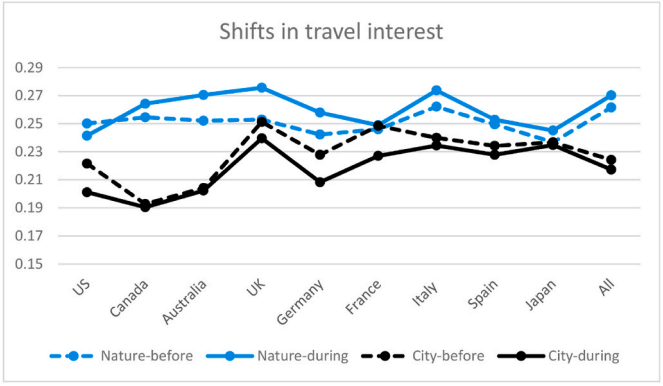


Fig. 3. Shift in travel interest (mean scores) for the selected destinations.

social capital, and thereby minimize their uncertainty and perceived risk in general (Rasoolimanesh et al., 2021). In particular, USA, UK and Canada could particularly be considered as the trustworthy countries because these countries are placed as the top countries in the 2019 Global Health Security Index for preparedness to manage a serious disease outbreak (Cameron, Nuzzo, & Bell, 2019; Mach et al., 2021). For all these countries, there was a decrease in city vs. nature tourism during Covid-19, which suggests a decrease in confidence in travel safety. The study by Mach et al. (2021) indicates that media coverage during Covid-19 created negative media imagery for many normally trustworthy destinations. This applied in particular to US media, where "patterns of U.*S. media* reporting were correlated with failures of national leadership under the Trump Administration, and they may have both reflected and contributed to politicization of Covid-19 in the United States" (Mach et al., 2021, p. 9). The negative destination image formation could possibly be explained by a combination of the framing effects of news media (Kapuściński & Richards, 2016), the naïve image of a destination (Cherifi et al., 2014; Jenkins, 1999; Luo & Zhai, 2017; Mason et al., 2005; Selby, 2004) and a generalization effect (Enders et al., 1992; Lepp & Gibson, 2008) developed through such media coverage (Yang et al., 2021). Tourism consumers might have generalized such negative naïve image of the US and other travel destinations, associating them with uncertainties and risks. Such naïve image and generalization effect might be applicable not only to a specific destination, but also to a specific traveling activity, such as "big city tourism". As consumers developed a naïve image about traveling to a big city, which can be associated with high risk of infection, they generalized big city tourism as a risk-bearing activity. Such naïve image and generalization of big city tourism might have affected some of the destinations typically considered, before Covid-19, as attractive big city destinations, such as New York, Berlin, Paris, and Melbourne (see Fig. 4) (Novelli et al., 2018; Schroeder & Pennington-Gray, 2014).

4.2. Study 2: sentiment and risk awareness – conservation vs. openness authors

Tables 7 and 8 respectively present mean scores of sentiment about traveling and risk awareness for conservation and openness authors during Covid-19. Fig. 5 and Table 9 further illustrate how travel sentiment and risk awareness of conservation and openness authors changed from before to during Covid-19.

5.3. Number of authors

Tables 7 and 8 both show that the number of conservation authors who posted during Covid-19 (1,471) is 150 lower than that of openness authors (1,621). Moreover, the distribution of conservation vs. openness authors varies by destination. The two groups are fairly balanced for USA and Canada, i.e., the number of authors is relatively high for both

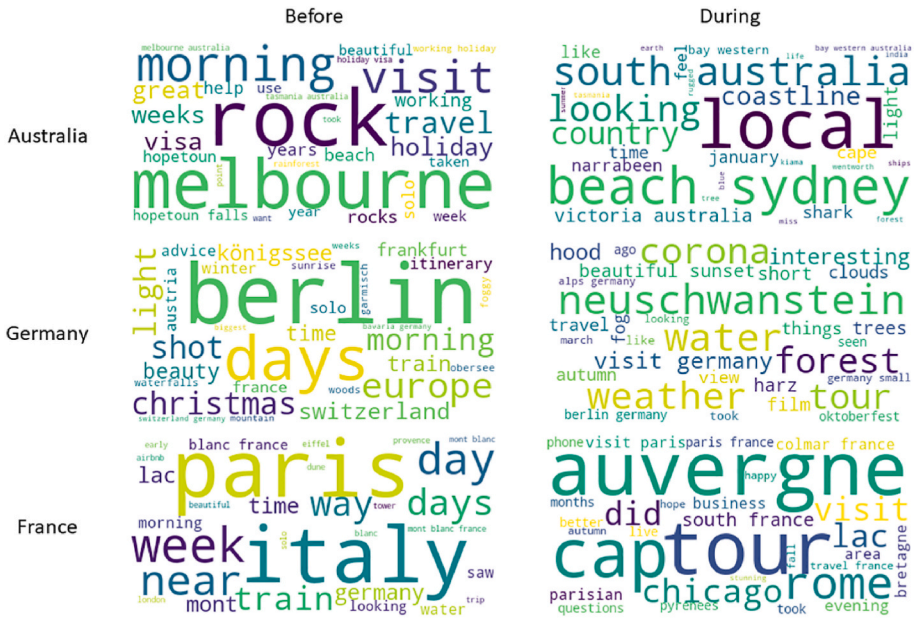


Fig. 4. Word clouds of posts mentioning the selected destinations Australia, Germany and France, before and during Covid-19.

Table 7
Comparison of travel *sentiment scores* between conservation vs. openness authors for selected destination during Covid-19.

	Conservation			Openness			Mean Differences		
	N	mean	SD	N	mean	SD	Dif.	*P*-value	Cohen's D
USA	90	0.097	0.328	100	0.120	0.342	−0.022	0.65	−0.066
Canada	70	0.048	0.262	58	0.121	0.328	−0.073	0.17	−0.248
Australia	33	0.165	0.303	42	0.098	0.322	0.068	0.36	0.215
UK	34	0.181	0.336	41	0.114	0.351	0.067	0.41	0.193
Germany	20	0.090	0.261	28	0.065	0.256	0.024	0.75	0.093
France	13	0.072	0.296	22	0.115	0.289	−0.043	0.68	−0.146
Italy	32	0.148	0.307	44	0.156	0.350	−0.008	0.92	−0.024
Spain	15	0.199	0.311	17	0.241	0.296	−0.042	0.70	−0.137
Japan	14	0.112	0.392	20	0.297	0.393	−0.185	0.19	−0.473
All During	1471	0.159	0.366	1621	0.210	0.398	−0.052	0.00***	−0.135
All Before	1471	0.269	0.450	1621	0.370	0.467	−0.101	0.00***	−0.220

Table 8
Comparison of *risk awareness scores* between conservation vs. openness authors for selected destination during Covid-19.

	Conservation			Openness			Mean Differences		
	N	mean	SD	N	mean	SD	Dif.	*P*-value	Cohen's D
USA	90	0.279	0.375	100	0.253	0.527	0.026	0.69	0.056
Canada	70	0.539	0.416	58	0.353	0.488	0.185	0.02**	0.412
Australia	33	0.778	0.541	42	0.616	0.435	0.162	0.17	0.335
UK	34	0.369	0.435	41	0.362	0.457	0.007	0.95	0.016
Germany	20	0.381	0.588	28	0.325	0.390	0.056	0.69	0.116
France	13	0.580	0.465	22	0.636	0.564	−0.057	0.75	−0.107
Italy	32	0.533	0.458	44	0.581	0.539	−0.048	0.68	−0.095
Spain	15	0.599	0.550	17	0.363	0.387	0.237	0.18	0.503
Japan	14	0.404	0.466	20	0.293	0.383	0.111	0.47	0.266
All During	1471	0.405	0.506	1621	0.364	0.521	0.041	0.03**	0.079
All Before	1471	0.382	0.459	1621	0.306	0.462	0.076	0.00***	0.165

Note1: p-values indicate statistical differences between conservation and openness authors ($p < 0.01$***, $p < 0.05$**, and $p < 0.1$*).
Note2: Cohen's D scores indicate effect size (0.2 refers to small effect size, 0.5 medium effect size, and 0.8 large effect size).
Note 3: Detailed statistics are found in Appendix 3.

conservation and openness authors for these destinations. On the other hand, the numbers of conservation authors mentioned about European destinations, Australia and Japan are lower compared to those of openness.

5.4. Sentiment of conservation vs. openness authors towards traveling

Authors' personal values of conservation and openness are determined based on their posts in the before1 period. Proposition 2b claims that authors with the openness value will have higher sentiment about traveling than the sentiment of the conservation authors. As shown in

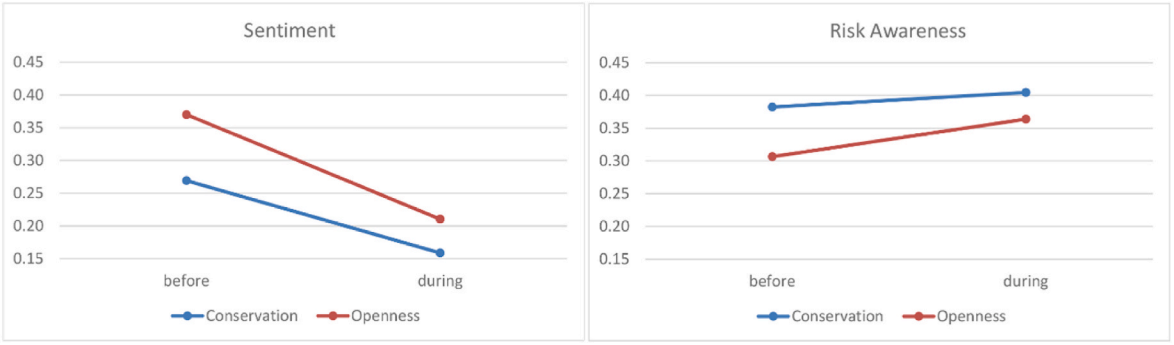


Fig. 5. Changes in travel sentiment and risk awareness by conservation (n = 1,471) and openness (n = 1,621) authors before and during Covid-19.

Table 9
Statistics of the changes in travel sentiment and risk awareness by conservation (n = 1,471) and openness (n = 1,621) authors before and during Covid-19.

	Mean before	Mean during	Paired Differences					t	df	Sig. (2-tailed)	Cohen's D
			Mean difference	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference					
						Lower	Upper				
Sentiment											
Conservation	0.2691	0.1586	0.11046	0.52184	0.01361	0.08377	0.13715	8.119	1470	0.000	0.212
Openness	0.3701	0.2103	0.15977	0.53634	0.01332	0.13364	0.18590	11.994	1620	0.000	0.298
Risk											
Conservation	0.3824	0.4045	−0.02211	0.60654	0.01581	−0.05313	0.00891	−1.398	1470	0.162	−0.036
Openness	0.3064	0.3638	−0.05743	0.60626	0.01506	−0.08696	−0.02789	−3.814	1620	0.000	−0.095

Note: Detailed statistics are found in Appendix 4a and Appendix 4b.

the bottom rows (all authors) of Table 7, openness authors have consistently higher sentiment about traveling than conservation authors in both before and during Covid-19. Specifically, openness authors have a mean travel sentiment of 0.37 and 0.21 respectively before and during Covid-19, while conservation authors have a mean travel sentiment of 0.26 and 0.16 respectively. While a large statistically significant difference in travel sentiment between conservation and openness is observed for both before and during Covid-19, the effect size in the difference becomes smaller in the during Covid-19 period. When looking at individual destinations, no statistical significance is observed.

Fig. 5 also demonstrates that Proposition 2b is supported for both before2 and during Covid-19. Moreover, the mean travel sentiments of both openness and conservation authors decreased significantly during Covid-19, which also supports Proposition 2a. A noteworthy observation is that the decrease of sentiment of travel-related Reddit posts is steeper for the openness authors. This trend is also confirmed in Table 9 showing that, while the statistical significance in the mean differences in travel sentiment before and during Covid-19 are observed for both types, the effect size (Cohen's D score) is larger for the openness authors.

5.5. Risk awareness

Fig. 5 and the bottom rows of Table 8 show that risk awareness of all openness authors is lower than that of all conservation authors both before and during Covid-19. Overall, conservation authors have a mean risk awareness of 0.38 and 0.40 before and during Covid-19 respectively, while openness authors have a lower risk awareness of 0.31 and 0.36. Although the size effects of these differences (Cohen's D) are small, these differences in risk awareness between conservation and openness are statistically significant, and supports our proposition 3b that openness authors have less risk awareness about travel than conservation authors do.

Fig. 5 and Table 9 also show that the mean risk awareness during Covid-19 increased significantly from before Covid-19 for the openness group. This further supports our proposition 3a. However, the increase of the mean risk awareness for the conservation group is not significant. The size effect of the decreases are very small in both openness and conservation authors. Table 8 shows that the risk awareness of openness authors during Covid-19 are generally lower than that of conservation authors in most of the selected destinations except France and Italy. However, statistically significant differences are not observed except for Canada.

5.6. Word clouds: conservation vs. openness authors

Word clouds in Fig. 6 show which words are frequently mentioned by authors before and during Covid-19. The first row compares word clouds before and during Covid-19 for all authors, whereas the second and third rows compare word clouds for conservation and openness, respectively.

In general, Covid-related words such as 'quarantine', 'refund', 'pandemic', and 'lockdown' are prominent during Covid-19, while words such as 'help', 'questions', 'suggestions', and 'recommendations' are prominent before Covid-19. The figure shows that openness authors tend to use words such as 'help', 'tips', 'question', and 'view', while prominent words for conservation authors include 'advice', 'solo', 'recommendation', and 'cruise' in the before Covid-19 period.

During Covid-19, the openness authors prominently use 'quarantine', 'pandemic', 'lockdown', 'stuck', and 'lucky, that may be a potential consequence of travel-related actions, while conservation authors used words that indicate future consideration of a trip, such as 'traveling', 'think', 'country', 'time year', 'spring', and 'October', during Covid-19.

5. Discussion

The results presented above confirm our two propositions about personal values: 2b) Sentiment about traveling expressed during Covid-19 by openness authors is more positive than that of conservation

Fig. 6. Word clouds expressed by all authors, conservation and openness authors before and during Covid-19.

authors; and 3b) Risk awareness of openness authors is lower than that of conservation authors. These differences are large and statistically significant.

The results also confirm our two other propositions: 2a) Sentiment about traveling during Covid-19 decreased from before Covid-19 for both openness and conservation authors; and 3a) Risk awareness increased during Covid-19 compared to before Covid-19 for both openness and conservation authors. However, the increase of the risk awareness is statistically significant only for the openness authors. In other words, our proposition 3a is not supported for the conservation authors. This could be related to the fact that we measured the risk awareness using the word embeddings of "risk" and "safe" without distinguishing tourism- or health-related risks (Zenker et al., 2021). As noted above, the number of conservation authors posting during Covid-19 is substantially lower than that of the openness authors in general. This tendency might reflect the fact that openness authors discuss more about actual traveling possibilities by considering potential travel-related risks (Zenker et al., 2021) as the word clouds of openness authors also indicate a potential consequence of travel-related actions, such as 'quarantine', 'stuck', and 'lucky'.

This suggests that many conservation authors during Covid-19 might be what Liu et al. (2016) categorize as high-risk/low-efficacy tourists; in other words, conservation authors are simply not considering traveling during this period, and therefore we do not observe a clear increase in their risk awareness as expressed by their posts. On the other hand, many openness authors might be categorized by Liu et al. as "proactive", or low-risk/high-efficacy. Although conscious of the risks, these authors are more likely to consider traveling and therefore post more about risks related to travel restrictions. This is consistent with the human value theory (Schwartz, 2006, 2007; Schwartz et al., 2012), which holds that openness individuals tend to be risk taking and sensation seeking (Zuckerman & Aluja, 2015), while conservation individuals are generally anxiety avoidant (Schwartz et al., 2012; Wolf et al., 2020). It is also consistent with the study by Kaptan et al. (2013), which reports that openness values are negatively correlated with risk from both emotional and cognitive perspectives, while conservation values are positively correlated with risk from an emotional perspective.

Finally, the statistically significant differences in the mean scores of sentiment and risk awareness are not observed for most individual destinations. This is a natural consequence of the smaller sample size for each destination. The number of conservation authors is higher than that of openness authors in Canada, whereas the number of conservation authors is particularly lower than that of openness in the European destinations, Australia and Japan. This tendency may in part reflect the fact that we have limited our data to English-language posts on the Reddit platform, thus perhaps resulting in more focus on English-speaking locations such as the U.S. Thus for many authors, the European and the ASEAN destinations might be seen as long-haul destinations. Therefore, conservation authors might have been less active in discussing about traveling far away in general during Covid-19.

On the other hand, the number of openness authors posted during Covid-19 is generally higher in these long-haul destinations. In addition, the results indicate some puzzling details: the mean scores of risk awareness for Italy and France by the openness authors are higher than those by the conservation authors. As suggested above, many openness authors could be classified as a low-risk/high efficacy proactive and sensation seeking segment (Liu et al., 2016; Zuckerman & Aluja, 2015) and have more experience in long-haul traveling (Floyd et al., 2004; Lepp & Gibson, 2003, 2008; Rasoolimanesh et al., 2021; Rittichainuwat & Chakraborty, 2009; Sönmez & Graefe, 1998) because they have generally higher sentiment about traveling and lower risk awareness in the before 2 period. Therefore, openness authors have discussed more about actual traveling possibilities by considering potential risks involving travel restrictions with popular long-haul destinations such as Italy and France. Accordingly, the significant increase of risk awareness observed for the openness authors during Covid-19 may be due to travel-related risks (Zenker et al., 2021) as potential travel risks typically appear when people start to talk about actually traveling to popular destinations such as Italy and France. As the European destinations such as Italy and France had country-specific Covid-19 measures, risks expected as the consequences of travel-related actions were diverse.

6.1. *Theoretical implications*

One of the foundational aspects of our study is the human value theory from Schwartz (1992) – a theory that has been called "the predominant model of values in psychology" (Wolf et al., 2020, p. 619). Notably, Schwartz (2012) defines values as a core "component of our self and personality" (p. 17), that is distinct from norms and attitudes. Work such as Sönmez and Graefe (1998) and Lepp and Gibson (2008) has shown that personality is important in determining risk aversion or risk seeking. Travel attitudes such as risk perception and travel-related

sentiment are also associated with the images individuals form of travel destinations (Baloglu & McCleary, 1999; Beerli & Martín, 2004). In the current work we have shown that values also play an important role here. We have shown that social media texts provide a sound basis for inferring personal values of individuals, and furthermore, these values can predict and explain subsequent travel attitudes and preferences, even in the chaotic and dynamic situation of a global pandemic.[5] In particular, we observe that individuals differ in their perception of risk as well as their general travel sentiment, based on whether they fall into the conservation or openness category of personal value.

Our study examines social media data over a period of several years, and we show that a classification of individuals' personal values in the first period is predictive of travel attitudes and preferences in two subsequent periods – the first prior to Covid-19, and the second during Covid-19. This long-term validity of personal value characteristics is consistent with the widely accepted view that personal values exhibit a "temporal stability" (Bilsky, Janik, & Schwartz, 2011, Wolf et al., 2020).

Glückstad et al. (2020) argue that destination image does not show the temporal stability of personal values; rather, it is continuously updated by stimuli. Naïve image formation of a travel destination has been argued to be sensitive to portrayals in the news media as well as information provided by governments and other institutions (Cherifi et al., 2014; Jenkins, 1999; Luo & Zhai, 2017; Mason et al., 2005; Selby, 2004). In a comparative study, Mach et al. (2021) show clear differences in media reporting along these dimensions in the selected countries (USA, Canada and UK) during Covid-19. Media reporting is widely understood to play a key role in destination image formation, especially during a crisis (Floyd et al., 2004; Lawton & Page, 1997; Villacé-Molinero et al., 2021). According to Rasoolimanesh et al. (2021) and Yang et al. (2021), media imagery was particularly important to individuals' formation of attitudes towards specific destinations and travel products during Covid-19. Kapuściński and Richards (2016) emphasize the framing effects of media coverage during pandemics and other emergencies, highlighting the moderating effects of psychographic characteristics. Similarly, Kaptan et al. (2013, p. 318) argue explicitly for "a link between personal values and risk perceptions".

The aforementioned literature provides a theoretical basis for incorporating personal values with the image formation and risk perception that underlies the empirical investigations reported here. A person who prioritizes a specific value (e.g., openness or conservation) relies on this value as a guiding principle in life (Sagiv, Roccas, Cieciuch, & Schwartz, 2017); thus personal values are relatively stable over time and closely connected to person's attitudes. On the other hand, image formation is based on a person's perception of an object - in the current case a travel destination. Therefore, image formation is dynamic, changing in response to external events and media portrayals. In our investigation, the dynamism of image formation and travel attitudes emerges forcefully during the Covid-19 pandemic; risk perception increases sharply, travel sentiment plummets, and there is a notable shift towards nature tourism away from city tourism. At the same time, we observe a stability in the way personal values affect individuals' response to these dynamic changes: both before and during Covid-19, openness individuals have higher travel sentiment and lower risk perception than conservation individuals.

These discussions suggest a potential extension of the existing theories. Our results strongly support the claim that personal values are stable over time, while image formation is more dynamic. A more fine-grained investigation of this would be fruitful: while image formation in general is dynamic, there might well be systematic variation in this. Some aspects of destination image might be relatively stable, while others change continually in response to changing conditions.

Schwartz's theory of values has an important cross-cultural perspective. In a study of twenty countries, Schwartz (2012) claims that the "circular structure, that captures the conflicts and compatibility among the ten values is apparently culturally universal (p. 2)". Similarly, Sagiv et al. (2017) indicate that the hierarchical structure of personal values reveals some commonality across cultures, whereas a "growing body of research indicates substantial variation in the importance attributed to values within and across cultures (p. 632)". In the current study we have not explicitly addressed issues of commonality or variation across cultures, but this would be an important future direction for research.

7. Methodological implications

Our study demonstrates the usefulness of NLP techniques to analyze social media data to explore travel-related preferences and attitudes. There are evident advantages in the use of social media data for such investigations. In particular, the use of social media makes it possible to collect a large amount of data in an open-ended format which individuals autonomously generated. There are also certain possible limitations in the use of social media – in our study, we relied on English-language posts on the Reddit platform. This means there would likely be an over-representation of authors from the North America. Furthermore, we did not have access to authors' socio-demographic information. Of course, any social media analysis is limited to authors who posted on a platform, which means that the samples used in the analysis may not be representative in terms of the overall population (Wang et al., 2019).

7.1. Managerial implications

From the managerial perspective, the fact that people are active on a travel-related social media platform is in itself an indication that they represent an attractive target segment for marketers. We have shown that recent advances in NLP technology make it possible to make inferences about the individuals within this attractive segment. Based on the text they produce, we are able to define subsegments of conservation and openness individuals, and we have shown clear differences between these subsegments in terms of travel-related sentiment and risk perception. This has evident relevance for marketing: tourism managers should target the openness authors who could be classified as a low-risk/high efficacy proactive segment (Liu et al., 2016), who have more experience in long-haul traveling (Floyd et al., 2004; Lepp & Gibson, 2003, 2008; Rasoolimanesh et al., 2021; Rittichainuwat & Chakraborty, 2009; Sönmez & Graefe, 1998). Since they have generally more positive sentiment about traveling and lower risk awareness in the before2 period, the communication should focus on their open-minded approach to experience, while avoiding travel-related risks.

Our study also shows how social media analysis can provide relevant qualitative insights in the form of word clouds. For example, Australia, France and Germany have words relevant to nature-oriented tourism being prominently mentioned during Covid-19. From the managerial perspective, these findings are valuable for designing a promotion for rural tourism arranged by a Destination Management Office (DMO). This indicates that the analytical design introduced in this article could be applicable for conducting a competitive analysis of multiple destinations in relation to their promotional activities.

Social media data can be used to complement survey-based approaches in consideration of its advantages and disadvantages. One limitation of our analysis is that we did not have access to authors' demographic profiles. An interesting future direction would be to perform an online survey targeting social media users. Such an online survey should integrate the original questions developed by Schwartz et al. (2012, 2017), which will enable us to classify social media users based on the actual survey consisting of other questions such as demographics, travel preferences, previous travel behavior and so on. The combination of the survey data with the social media classification techniques presented in this study will enhance the quality of the social media analytical framework, as is suggested by the work of Chen et al. (2014). Such a survey would produce data that could be used to train an NLP system to make much more fine-grained inferences concerning

travel preferences and attitudes. Given the continuing improvements in NLP technology, in particular large language models (Brown et al., 2020), we envision future work in which marketing managers can perform very fine-grained segmentation or even individualized targeting of potential travelers, based on text found on social media.

8. Conclusion

In a crisis situation, tourism managers need to identify potentially attractive market segments consisting of individuals who will consider traveling to a destination despite potential risks of health-related and travel-related issues (Dolnicar, 2005, 2006; Ritchie & Jiang, 2019). It is also important for tourism managers to gain insight into the way potential visitors perceive different tourism products and destinations during and after a crisis (Hajibaba et al., 2015). The current study uses social media data to explore consumers' preferences for city vs. nature products, and it also classifies consumers' personal values to explore how these values relate to general attitudes relating to travel. We collected over a million travel-related posts, in the first large-scale examination of social media data as a reflection of personal values, travel-oriented attitudes and preferences. This study crucially relies on recently developed techniques to perform semantic measurements using word embeddings and the measurement of vector distances. This is what makes it possible to determine the values of conservation and openness in authors' texts, and also to measure their risk awareness, travel sentiment, and preferences for city vs. nature tourism.

The analysis demonstrated that interest in nature tourism increased during Covid-19 in general, but especially for resource-rich destinations such as UK, Canada, Germany and Australia; interest in city tourism declined, in particular for destinations with big cities such as the US (e.g. NY), France (e.g. Paris) and Germany (e.g. Berlin). These findings are consistent with existing research such as Huang et al. (2021); Lee and Lee (2020); Marques et al. (2021) and can be explained by people's desire for social distancing during a health-related crisis.

It has been argued that personal values of consumers are relevant to understanding their preferences and attitudes, concerning travel as well as other domains (Schwartz, 2006, 2007; Schwartz et al., 2012, 2017). While these claims have received some support based on survey data, the present study provides a new perspective on the association of consumers with personal values. By using a large collection of social media posts, we were able to apply NLP techniques to classify consumers based on these personal values, and furthermore, we showed that these are highly predictive of their travel-related attitudes, both before and during Covid-19. Our analysis of multiple indicators – i.e., numbers of authors and posts, travel-related sentiment and risk awareness concerning different destinations, as well as the general shift in travel-related sentiment and risk awareness before and during Covid-19 – signpost distinctive patterns associated with characteristics of conservation and openness authors; segments that have been defined based on posts prior to Covid-19. These patterns have clear managerial implications, as well as theoretical interest, and the results reported here suggest that the use of NLP to analyze social media data represents a promising direction for continuing research in the attitudes and preferences of travelers.

Impact statement

This article applies natural language processing techniques (NLP) to analyze over one million travel-related posts on the social media platform Reddit, collected from January 2016 to January 2021 on travel-related subreddits. Our study demonstrates the usefulness of NLP to classify consumers based on posts before Covid-19 as well as its predictiveness of key travel-related attributes both before and during Covid-19: namely, risk-awareness and sentiment about traveling. The study further demonstrates that nature tourism gained overall during Covid-19 when compared to city tourism, particularly in destinations with rich natural resources. From a managerial perspective, these findings are valuable for designing promotions of specific types of tourism, targeting segments identified by social media analysis. This study also shows how social media analysis based on NLP can provide insight into the applicability of psychographic theories and the postulation of personal values, as sources of insight into the preferences and attitudes of travelers. It also shows that newly developed methods of semantic analysis mean that social media analysis can be an interesting alternative to survey-based methods.

Credit author statement

Daniel Hardt: Conceptualization: Methodology: Software: Resources: Data curation: Investigation: Visualization: Writing – Original Draft: Writing – Review & Editing.

Fumiko Kano Glückstad: Conceptualization: Methodology: Investigation: Formal analysis: Visualization: Writing – Original Draft: Writing – Review & Editing.

Funding source

None.

Declaration of competing interest

The authors report no conflicts of interest.

Acknowledgement

We appreciate the anonymous reviewers who provided highly constructive and encouraging suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tourman.2023.104821.

References

Abubakar, A. M., Ilkan, M., Meshall Al-Tal, R., & Eluwole, K. K. (2017). eWOM, revisit intention, destination trust and gender. *Journal of Hospitality and Tourism Management, 31*, 220–227. https://doi.org/10.1016/j.jhtm.2016.12.005. Retrieved from doi: 10.1016/j.jhtm.2016.12.005.

Almeida, F., & Xexéo, G. (2019). *Word embeddings: A survey*. arXiv preprint arXiv: 1901.09069.

Google Code Archive. (2013). Word2vec. https://code.google.com/archive/p/word2vec/.

Arefieva, V., Egger, R., & Yu, J. (2021). A machine learning approach to cluster destination image on Instagram. *Tourism Management, 85*(March), Article 104318. https://doi.org/10.1016/j.tourman.2021.104318. Retrieved from doi: 10.1016/j.tourman.2021.104318.

Backer, E., & Ritchie, B. W. (2017). VFR travel: A viable market for tourism crisis and disaster recovery? *International Journal of Tourism Research, 19*(4), 400–411. https://doi.org/10.1002/jtr.2102

Baloglu, S., & McCleary, K. W. (1999). A model of destination image formation. *Annals of Tourism Research, 26*(4), 868–897. https://doi.org/10.1016/S0160-7383(99)00030-4

Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin, 29*(10), 1207–1220. https://doi.org/10.1177/0146167203254602

Beerli, A., & Martín, J. D. (2004). Factors influencing destination image. *Annals of Tourism Research, 31*(3), 657–681. https://doi.org/10.1016/j.annals.2004.01.010

Beirman, D. (2002). Marketing of tourism destinations during a prolongedcrisis: Israel and the Middle East. *Journal of Vacation Marketing, 8*(2), 167–176. https://doi.org/10.1177/135676670200800206

Bilsky, W., Janik, M., & Schwartz, S. H. (2011). The structural organiza- tion of human values: Evidence from three rounds of the European Social Survey (ESS). *Journal of Cross-Cultural Psychology, 42*, 759–776. https://doi.org/10.1177/0022022110362757

Bjørkelund, E., Burnett, T. H., & Nørvåg, K. (2012). A study of opinion mining and visualization of hotel reviews. *ACM International Conference Proceeding Series*, 229–238. https://doi.org/10.1145/2428736.2428773

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.

Cahyanto, I., Wiblishauser, M., Pennington-Gray, L., & Schroeder, A. (2016). *The dynamics of travel avoidance: The case of Ebola in the U.S. Tourism Management Perspectives, 20*, 195–203. https://doi.org/10.1016/j.tmp.2016.09.004. Retrieved from.

Cameron, E. E., Nuzzo, J. B., & Bell, J. A. (2019). *GHS index: Global health security index: Building collective action and accountability*. NTI and Johns Hopkins Bloomberg School of Public Health.

Cavlek, N. (2002). Tour operators and destination safety. *Annals of Tourism Research, 29*(2), 478–496. https://doi.org/10.1016/S0160-7383(01)00067-6

Chen, J., Hsieh, G., Mahmud, J., & Nichols, J. (2014). Understanding individuals' personal values from social media word use. Proceedings of the ACM Conference on Computer Supported CooperativeWork. *Computer Supported Cooperative Work, 405*. https://doi.org/10.1145/2531602.2531608. –414.

Cherifi, B., Smith, A., Maitland, R., & Stevenson, N. (2014). Destination images of non-visitors. *Annals of Tourism Research, 49*, 190–202. https://doi.org/10.1016/j.annals.2014.09.008

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., … Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science, 368*(6489), 395–400. https://doi.org/10.1126/science.aba9757

Claster, W. B., Cooper, M., & Sallis, P. (2010). Thailand - tourism and conflict. Modeling sentiment from twitter tweets using naïve bayes and unsupervised artificial neural nets. *Proceedings - 2nd International Conference on Computational Intelligence, Modelling and Simulation, CIMSim, 2010*, 89–94. https://doi.org/10.1109/CIMSiM.2010.98

Claster, W. B., Dinh, H., & Cooper, M. (2010). Naive bayes and unsupervised artificial neural nets for Cancun tourism social media data analysis. In *Proceedings - 2010 2nd world congress on nature and biologically inspired computing* (pp. 158–163). NaBIC. https://doi.org/10.1109/NABIC.2010.5716370.

Collins-Kreiner, N., & Ram, Y. (2020). *National tourism strategies during the Covid-19 pandemic*. January: Annals of Tourism Research.

Culligan, K., Dubber, J., & Lotten, M. (2014). *As others see us: Culture, attraction and soft power (tech. Rep.)*. British Council.

Dolnicar, S. (2005). Understanding barriers to leisure travel: Tourist fears as a marketing basis. *Journal of Vacation Marketing, 11*(3), 197–208. https://doi.org/10.1177/1356766705055706

Dolnicar, S. (2006). Crises that scare tourists: Investigating tourists travel-related concerns. *Crisis Management in Tourism*, 98–109. https://doi.org/10.1079/9781845930479.0098

Elsaadawy, A., Torki, M., & Ei-Makky, N. (2018). A text classifier using weighted average word embedding. In *2018 international Japan africa conference on electronics, communications and computations (jac-ecc)* (pp. 151–154).

Enders, W., Sandler, T., & Parise, G. F. (1992). An econometric analysis of the impact of terrorism on tourism. *Kyklos, 45*(4), 531–554. https://doi.org/10.1111/j.1467-6435.1992.tb02758.x

Fischer, R., & Smith, P. B. (2006). Who cares about justice? The moderating effect of values on the link between organisational justice and work behaviour. *Applied Psychology, 55*(4), 541–562. https://doi.org/10.1111/j.1464-0597.2006.00243.x

Floyd, M. F., Gibson, H., Pennington-Gray, L., & Thapa, B. (2004). The effect of risk perceptions on intentions to travel in the aftermath of september 11, 2001. *Journal of Travel & Tourism Marketing, 15*(2–3), 19–38. https://doi.org/10.1300/J073v15n02_02

Gkritzali, A. (2017). Online destination sentiment recovery during a sustained crisis. *Annals of Tourism Research, 66*, 183–185. https://doi.org/10.1016/j.annals.2017.05.012. Retrieved from doi:10.1016/j.annals.2017.05.012.

Glückstad, F. K., Schmidt, M. N., & Mørup, M. (2020). Testing a model of destination image formation: Application of Bayesian relational modelling and fsQCA. *Journal of Business Research, 120*(October 2019), 351–363. https://doi.org/10.1016/j.jbusres.2019.10.014

Glückstad, F. K., Wiil, U. K., Mansourvar, M., & Andersen, P. T. (2021). Cross-cultural bayesian network analysis of factors affecting residents' concerns about the spread of an infectious disease caused by tourism. *Frontiers in Psychology, 12*(June), 1–19. https://doi.org/10.3389/fpsyg.2021.635110

Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: A rapid assessment of COVID-19. *Journal of Sustainable Tourism, 29*(1), 1–20. https://doi.org/10.1080/09669582.2020.1758708. Retrieved from doi: 10.1080/09669582.2020.1758708.

Govers, R., Go, F. M., & Kumar, K. (2007). Virtual destination image a new measurement approach. *Annals of Tourism Research, 34*(4), 977–997. https://doi.org/10.1016/j.annals.2007.06.001

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. Retrieved from doi: 10.1016/j.tourman.2016.09.009.

Hajibaba, H., Gretzel, U., Leisch, F., & Dolnicar, S. (2015). Crisis-resistant tourists. *Annals of Tourism Research, 53*, 46–60. https://doi.org/10.1016/j.annals.2015.04.001. Retrieved from doi: 10.1016/j.annals.2015.04.001.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., et al. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour, 5*(4), 529–538. https://doi.org/10.1038/s41562-021-01079-8

Handler, I. (2016). The impact of the Fukushima disaster on Japan's travel image: An exploratory study on Taiwanese travellers. *Journal of Hospitality and Tourism Management, 27*, 12–17. https://doi.org/10.1016/j.jhtm.2016.01.007. Retrieved from doi: 10.1016/j.jhtm.2016.01.007.

Han, Q., Zheng, B., Cristea, M., Agostini, M., Belanger, J. J., Gutzkow, B., et al. (2021). Trust in government regarding COVID-19 and its associations with preventive health behavior and prosocial behaviour during the pandemic: A cross-sectional and longitudinal study. *Psychological Medicine*. https://doi.org/10.1017/S0033291721001306

Henderson, J. C. (2007). *Tourism crises: Causes, consequences and management*. Elsevier Inc.

Henderson, J. C., & Ng, A. (2004). Responding to crisis: Severe acute respiratory syndrome (SARS) and hotels in Singapore. *International Journal of Tourism Research, 6*(6), 411–419. https://doi.org/10.1002/jtr.505

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. SAGE Publications. Retrieved from https://books.google.dk/books?id=hW6AAAAMAAJ.

Huang, S. S., Shao, Y., Zeng, Y., Liu, X., & Li, Z. (2021). Impacts of COVID-19 on Chinese nationals' tourism preferences. *Tourism Management Perspectives, 40*(July), Article 100895. https://doi.org/10.1016/j.tmp.2021.100895. Retrieved from doi: 10.1016/j.tmp.2021.100895.

Hu, Y., & Ritchie, J. B. (1993). Measuring destination attractiveness: A contextual approach. *Journal of Travel Research, 32*(2), 25–34. https://doi.org/10.1177/004728759303200204

Im, J., Kim, J., & Choeh, J. Y. (2021). COVID-19, social distancing, and risk-averse actions of hospitality and tourism consumers: A case of South Korea. *Journal of Destination Marketing & Management, 20*, Article 100566. https://doi.org/10.1016/j.jdmm.2021.100566. Retrieved from doi: 10.1016/j.jdmm.2021.100566.

Jenkins, O. (1999). Understanding and measuring tourist destination images. *International Journal of Tourism Research, 1*, 1–15.

Kaptan, G., Shiloh, S., & Önkal, D. (2013). Values and risk perceptions: A cross-cultural examination. *Risk Analysis, 33*(2), 318–332. https://doi.org/10.1111/j.1539-6924.2012.01875.x

Kapuściński, G., & Richards, B. (2016). News framing effects on destination risk perception. *Tourism Management, 57*, 234–244. https://doi.org/10.1016/j.tourman.2016.06.017

Kozak, M., Crotts, J. C., & Law, R. (2007). The impact of the perception of risk on international travellers. *International Journal of Tourism Research, 9*(4), 233–242. https://doi.org/10.1002/jtr.607

Lawton, G., & Page, S. (1997). Evaluating travel agents' provision of health advice to travellers. *Tourism Management, 18*(2), 89–104. https://doi.org/10.1016/S0261-5177(96)00104-5

Lee, S. M., & Lee, D. H. (2020). *"Untact": A New Customer Service Strategy in the Digital Age, 14*(No. 1). https://doi.org/10.1007/s11628-019-00408-2

Lepp, A., & Gibson, H. (2003). Tourist roles, perceived risk and international tourism. *Annals of Tourism Research, 30*(3), 606–624. https://doi.org/10.1016/S0160-7383(03)00024-0

Lepp, A., & Gibson, H. (2008). Sensation seeking and tourism: Tourist role, perception of risk and destination choice. *Tourism Management, 29*(4), 740–750. https://doi.org/10.1016/j.tourman.2007.08.002

Levy, O., Goldberg, Y., & Ido, D. (2015). Improving distributional similarity with lessons learned from word embeddings. *Transactions of the association for computational linguistics3*, 211–225.

Li, Y. R., Lin, Y. C., Tsai, P. H., & Wang, Y. Y. (2015). Traveller-generated contents for destination image formation: Mainland China travellers to taiwan as a case study. *Journal of Travel & Tourism Marketing, 32*(5), 518–533. https://doi.org/10.1080/10548408.2014.918924

Li, J., Nguyen, T. H. H., & Coca-Stefaniak, J. A. (2021). *Coronavirus impacts on post-pandemic planned travel behaviours* (Vol. 86). https://doi.org/10.1016/j.annals.2020.102964

Liu, B., Schroeder, A., Pennington-Gray, L., & Farajat, S. A. (2016). Source market perceptions: How risky is Jordan to travel to? *Journal of Destination Marketing & Management, 5*(4), 294–304. https://doi.org/10.1016/j.jdmm.2016.08.005. Retrieved from doi: 10.1016/j.jdmm.2016.08.005.

Luo, Q., & Zhai, X. (2017). "I will never go to Hong Kong again!" How the secondary crisis communication of "Occupy Central" on Weibo shifted to a tourism boycott. *Tourism Management, 62*, 159–172. https://doi.org/10.1016/j.tourman.2017.04.007. Retrieved from doi: 10.1016/j.tourman.2017.04.007.

Mach, K. J., Salas Reyes, R., Pentz, B., Taylor, J., Costa, C. A., Cruz, S. G., et al. (2021). News media coverage of COVID-19 public health and policy information. *Humanities and Social Sciences Communications, 8*(1), 1–11. https://doi.org/10.1057/s41599-021-00900-z

Marques, C. P., Guedes, A., & Bento, R. (2021). *Rural tourism recovery between two COVID-19 waves: The case of Portugal*. https://doi.org/10.1080/13683500.2021.1910216

Mason, P., Grabowski, P., & Du, W. (2005). Severe acute respiratory syndrome, tourism and the media. *International Journal of Tourism Research, 7*(1), 11–21. https://doi.org/10.1002/jtr.519

Mazanec, J. A., & Strasser, H. (2007). Perceptions-based analysis of tourism products and service providers. *Journal of Travel Research, 45*(4), 387–401. https://doi.org/10.1177/0047287507299576

McKercher, B., & Chon, K. (2004). The over-reaction to SARS and the collapse of Asian tourism. *Annals of Tourism Research, 31*(3), 716–719. https://doi.org/10.1016/j.annals.2003.11.002

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. arXiv preprint arXiv:1301.3781.

Mizrachi, I., & Fuchs, G. (2016). Should we cancel? An examination of risk handling in travel social media before visiting ebola-free destinations. *Journal of Hospitality and Tourism Management, 28*, 59–65. https://doi.org/10.1016/j.jhtm.2016.01.009. Retrieved fromdoi: 10.1016/j.jhtm.2016.01.009.

Möller, C., Wang, J., & Nguyen, H. T. (n.d.). doi: 10.1016/j.tourman.2018.05.014.

Neuburger, L., & Egger, R. (2021). Travel risk perception and travel behavior during the COVID-19 pandemic 2020: A case study of the DACH region. *Current Issues in Tourism, 24*(7), 1003–1016. https://doi.org/10.1080/13683500.2020.1803807. Retrieved from doi:10.1080/13683500.2020.1803807.

Novelli, M., Gussing Burgess, L., Jones, A., & Ritchie, B. W. (2018). 'No Ebola. . . still doomed' – the Ebola-induced tourism crisis. *Annals of Tourism Research, 70*(March), 76–87. https://doi.org/10.1016/j.annals.2018.03.006. Retrieved from.

Orkphol, K., & Wu, Y. (2019). Word sense disambiguation using cosine similarity collaborates with Word2vec and WordNet. *Future Internet, 5*, 114.

Pantano, E., Priporas, C. V., Devereux, L., & Pizzi, G. (2021). *Tweets to escape: Intercultural differences in consumer expectations and risk behavior during the COVID-19 lockdown in three European countries* (Vol. 130). https://doi.org/10.1016/j.jbusres.2021.03.015

Parady, G., Taniguchi, A., & Takami, K. (2020). Travel behavior changes during the COVID-19 pandemic in Japan: Analyzing the effects of risk perception and social influence on going-out self-restriction. *Transportation Research Interdisciplinary Perspectives, 7*, Article 100181. https://doi.org/10.1016/j.trip.2020.100181. Retrieved from doi:10.1016/j.trip.2020.100181.

Rasoolimanesh, S. M., Seyfi, S., Rastegar, R., & Hall, C. M. (2021). Destination image during the COVID-19 pandemic and future travel behavior: The moderating role of past experience. *Journal of Destination Marketing & Management, 21*(May), Article 100620. https://doi.org/10.1016/j.jdmm.2021.100620. Retrieved from doi: 10.1016/j.jdmm.2021.100620.

Reichel, A., Fuchs, G., & Uriely, N. (2007). Perceived risk and the non-institutionalized tourist role: The case of Israeli student exbackpackers. *Journal of Travel Research, 46*(2), 217–226. https://doi.org/10.1177/0047287507299580

Reisinger, Y., & Mavondo, F. (2005). Travel anxiety and intentions to travel internationally: Implications of travel risk perception. *Journal of Travel Research, 43*(3), 212–225. https://doi.org/10.1177/0047287504272017

Ritchie, B. W., Crotts, J. C., Zehrer, A., & Volsky, G. T. (2014). Understanding the effects of a tourism crisis: The impact of the BP oil spill on regional lodging demand. *Journal of Travel Research, 53*(1), 12–25. https://doi.org/10.1177/0047287513482775

Ritchie, B. W., & Jiang, Y. (2019). A review of research on tourism risk, crisis and disaster management: Launching the annals of tourism research curated collection on tourism risk, crisis and disaster management. *Annals of Tourism Research, 79*(November), Article 102812. https://doi.org/10.1016/j.annals.2019.102812. Retrieved from doi:10.1016/j.annals.2019.102812.

Rittichainuwat, B. N., & Chakraborty, G. (2009). Perceived travel risks regarding terrorism and disease: The case of Thailand. *Tourism Management, 30*(3), 410–418. https://doi.org/10.1016/j.tourman.2008.08.001

Sagiv, L., Roccas, S., Cieciuch, J., & Schwartz, S. H. (2017). Personal values in human life. *Nature Human Behaviour, 1*(September). https://doi.org/10.1038/s41562-017-0185-3

Sánchez-Cañizares, S. M., Cabeza-Ramírez, L. J., Muñoz-Fernández, G., & Fuentes-García, F. J. (2021). Impact of the perceived risk from Covid-19 on intention to travel. *Current Issues in Tourism, 24*(7), 970–984. https://doi.org/10.1080/13683500.2020.1829571. Retrieved from doi: 10.1080/13683500.2020.1829571.

Santos, M., Madrid González, A., Haegeman, C., & Rainoldi, K. (2020). Behavioural changes in tourism in times of COVID-19 Employment scenarios and policy options. In *Public details (issue august*. https://doi.org/10.2760/00411. Jouint Research Center (European Commission).

Schroeder, A., & Pennington-Gray, L. (2014). Perceptions of crime at the Olympic Games: What role does media, travel advisories, and social media play? *Journal of Vacation Marketing, 20*(3), 225–237. https://doi.org/10.1177/1356766714521809

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25*(C), 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H. (2006). *Basic human values: An overview* (Vol. 48, pp. 49–71). Jerusalem Hebrew University. https://doi.org/10.1080/026999499377664

Schwartz, S. H. (2007). Value orientations: Measurement, antecedents and consequences across nations. *Measuring attitudes cross-nationally: Lessons from the European Social Survey*, 169–203.

Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture, 2*(1), 1–20. https://doi.org/10.9707/2307-0919.1116

Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., … Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology, 103*(4), 663–688. https://doi.org/10.1037/a0029393

Schwartz, S. H., Cieciuch, J., Vecchione, M., Torres, C., Dirilen-Gumus, O., & Butenko, T. (2017). Value tradeoffs propel and inhibit behavior: Validating the 19 refined values in four countries. *European Journal of Social Psychology, 47*(3), 241–258. https://doi.org/10.1002/ejsp.2228

Seabra, C., Dolnicar, S., Abrantes, J. L., & Kastenholz, E. (2013). Heterogeneity in risk and safety perceptions of international tourists. *Tourism Management, 36*(1), 502–510. https://doi.org/10.1016/j.tourman.2012.09.008. Retrieved from doi: 10.1016/j.tourman.2012.09.008.

Selby, M. (2004). *Understanding urban tourism: Image culture & experience*. NY: Ltd., I.B. auris & Co Ltd.

Sharma, A., & Daniels, A. (2020). *"Tweets Sentiment Analysis via Word Embeddings and Machine Learning Techniques."arXiv preprint arXiv, 2007*, Article 04303.

Shiina, A., Niitsu, T., Kobori, O., Idemoto, K., Hashimoto, T., Sasaki, T., & Iyo, M. (2020). Relationship between perception and anxiety about COVID-19 infection and risk behaviors for spreading infection: A national survey in Japan. *Brain, Behavior, & Immunity - Health, 6*(July), Article 100101. https://doi.org/10.1016/j.bbih.2020.100101. Retrieved fromdoi: 10.1016/j.bbih.2020.100101.

Shimada, K., Inoue, S., Maeda, H., & Endo, T. (2011). Analyzing tourism information on twitter for a local city. *Proceedings - 1st ACIS International Symposium on Software and Network Engineering, SSNE, 2*(2011), 61–66. https://doi.org/10.1109/SSNE.2011.27

Sönmez, S. F., Apostolopoulos, Y., & Tarlow, P. (1999). Tourism in crisis: Managing the effects of terrorism. *Journal of Travel Research, 38*(1), 13–18. https://doi.org/10.1177/004728759903800104

Sönmez, S. F., & Graefe, A. R. (1998). Influence of terrorism risk on foreign tourism decisions. *Annals of Tourism Research, 25*(1), 112–144. https://doi.org/10.1016/s01607383(97)00072-8

Taylor, T., & Toohey, K. (2007). Perceptions of terrorism threats at the 2004 olympic games: Implications for sport events. *Journal of Sport & Tourism, 12*(2), 99–114. https://doi.org/10.1080/14775080701654754

Team, N. (2021). *Natural Language Toolkit*. https://www.nltk.org.

Vaishar, A., & Šťastná, M. (2022). Impact of the COVID-19 pandemic on rural tourism in Czechia Preliminary considerations. *Current Issues in Tourism, 25*(Issue 2), 187–191. https://doi.org/10.1080/13683500.2020.1839027

Villacé-Molinero, T., Fernández-Muñoz, J. J., Orea-Giner, A., & Fuentes-Moraleda, L. (2021). Understanding the new post-COVID-19 risk scenario: Outlooks and challenges for a new era of tourism. *Tourism Management, 86*(March). https://doi.org/10.1016/j.tourman.2021.104324

Vu, H. Q., Li, G., & Law, R. (2020). Cross-country analysis of tourist activities based on venue-referenced social media data. *Journal of Travel Research, 59*(1), 90–106. https://doi.org/10.1177/0047287518820194

Wachyuni, S. S., & Kusumaningrum, D. A. (2020). The effect of COVID-19 pandemic: How are the future tourist behavior? *Journal of Education, Society and Behavioural Science, August*, 67–76. https://doi.org/10.9734/jesbs/2020/v33i430219

Wang, Z., Hale, S., Adelani, D. I., Grabowicz, P., Hartman, T., Flöck, F., et al. (2019). Demographic inference and representative population estimates from multilingual social media data. In *The world wide web conference* (pp. 2056–2067).

Wang, J., & Ritchie, B. W. (2012). Understanding accommodation managers' crisis planning intention: An application of the theory of planned behaviour. *Tourism Management, 33*(5), 1057–1067. https://doi.org/10.1016/j.tourman.2011.12.006. Retrieved from doi:10.1016/j.tourman.2011.12.006.

Washer, P. (2004). Representations of SARS in the British newspapers. *Social Science & Medicine, 59*(12), 2561–2571. https://doi.org/10.1016/j.socscimed.2004.03.038

Wolff, K., Larsen, S., & Øgaard, T. (2019). How to define and measure risk perceptions. *Annals of Tourism Research, 79*(August). https://doi.org/10.1016/j.annals.2019.102759

Wolf, L. J., Haddock, G., Manstead, A. S., & Maio, G. R. (2020). The importance of (shared) human values for containing the COVID-19 pandemic. *British Journal of Social Psychology, 59*(3), 618–627. https://doi.org/10.1111/bjso.12401

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015a). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management, 44*, 120–130. https://doi.org/10.1016/j.ijhm.2014.10.013

Xiang, Z., Schwartz, Z., & Uysal, M. (2015b). What types of hotels make their guests (Un) Happy? Text analytics of customer experiences in online reviews. *Information and Communication Technologies in Tourism, 2015*, 33–45. https://doi.org/10.1007/978-3-319-14343-9_3

Yang, S., Isa, S. M., & Ramayah, T. (2021). How are destination image and travel intention influenced by misleading media coverage? Consequences of COVID-19 outbreak in China. *Vision*. https://doi.org/10.1177/0972262921993245

Zenker, S., Braun, E., & Gyimóthy, S. (2021). Too afraid to travel? Development of a pandemic (COVID-19) anxiety travel scale (PATS). *Tourism Management, 84*(December 2020). https://doi.org/10.1016/j.tourman.2021.104286

Zheng, D., Luo, Q., & Ritchie, B. W. (2021). Afraid to travel after COVID-19? Self-Protection, coping and resilience against pandemic 'travel fear'. *Tourism Management, 83*(October 2020), Article 104261. https://doi.org/10.1016/j.tourman.2020.104261. Retrieved fromdoi:10.1016/j.tourman.2020.104261.

Zuckerman, M., & Aluja, A. (2015). Measures of Sensation Seeking. In *Measures of Personality and Social Psychological Constructs*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-386915-9.00013-9.

Daniel Hardt is Associate Professor in computational linguistics at Copenhagen Business School. His research involves natural language processing, theoretical linguistics, AI and machine learning. He has over 60 peer-reviewed publications, and his work has appeared in Computational Linguistics, Linguistics and Philosophy, Journal of Semantics, Big Data, and The European Journal of Information Systems.Contribution to the manuscript: conceptualization of the study, overall study design, data collection, data mining, NLP analysis of data, visualization, manuscript draft.

Fumiko Kano Glückstad is Associate Professor in intercultural market communication at Copenhagen Business School. Her research interests are cross-cultural psychology, consumer psychology, typological analysis, and profiling of consumer segments. She has experience in an interdisciplinary tourism research project that investigated consumers' mental destination representation and consumer segmentation. Her publications appear in Journal of Business Research, Journal of Cross-Cultural Psychology, Journal of Mathematical Sociology among others. Contribution to the manuscript: conceptualization of the study, overall study design, data analysis and interpretation of results, manuscript draft.