# A systematic literature review on the use of big data for sustainable tourism

Rahmadian, Eko; Feitosa, Daniel; Zwitter, Andrej

*Published in:*
Current Issues in Tourism

*DOI:*
10.1080/13683500.2021.1974358



*Publication date:*
2022

Routledge
Taylor & Francis Group

Current Issues in Tourism

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rcit20

# A systematic literature review on the use of big data for sustainable tourism

E. Rahmadian, D. Feitosa & A. Zwitter

**To cite this article:** E. Rahmadian, D. Feitosa & A. Zwitter (2022) A systematic literature review on the use of big data for sustainable tourism, Current Issues in Tourism, 25:11, 1711-1730, DOI: 10.1080/13683500.2021.1974358

**To link to this article:** https://doi.org/10.1080/13683500.2021.1974358



Published online: 14 Sep 2021.

CURRENT ISSUES IN TOURISM
2022, VOL. 25, NO. 11, 1711–1730
https://doi.org/10.1080/13683500.2021.1974358

Routledge
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

# A systematic literature review on the use of big data for sustainable tourism

E. Rahmadian, D. Feitosa and A. Zwitter

Campus Fryslan, University of Groningen, Groningen, Netherlands

**ABSTRACT**

Sustainable tourism research focuses on mitigating or remediating environmental, social and economic impacts on tourism. In the past years, Big Data approaches have been applied to the field of tourism allowing for remarkable progress. However, there seems to be little evidence to support that such approaches are an inspiration to sustainable tourism and are being implemented. In this context, we aim to obtain a comprehensive overview of the use of Big Data in sustainable tourism to address various issues and understand how Big Data can support decision-making in such scenarios. To that end, this paper reports on the results of a literature review via a combination of a Systematic Literature Review (SLR) in Software Engineering, and the use of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method. In summary, we investigated four facets: (a) sources of big data, (b) approaches, (c) purposes, and (d) contexts of application. The results suggest that the use of various approaches have impacted practices in sustainable tourism. The findings provide a thorough understanding of the state of the art of Big Data application in sustainable tourism and provide valuable insights to foster growth both in terms of research and practice.

**ARTICLE HISTORY**
Received 18 March 2021
Accepted 24 August 2021

**KEYWORDS**
Sustainable tourism; big data; systematic literature review; Internet of Things; artificial intelligence

## 1. Introduction

Big data has attracted substantial interest from researchers as a means to obtain insights to find solutions, new strategies, and to uncover hidden potentials for an array of purposes (Mayer-Schonberger & Cukier, 2013). Big data is a term that primarily describes datasets that are so large, unstructured and complex that they require advanced and unique technologies to store, manage, analyse and visualize data (Gunther et al., 2017). Datasets with such features are emerging from multiple economic sectors. Among them, the tourism sector has distinctive characteristics, i.e. the intangibility and perishability of products and the vulnerability to natural, economic and political phenomena.

The tourism industry thrives on information. As such, research on tourism requires large volumes of vigorous, updated, timely and relevant information to support and help decision-making processes. Big data can deliver up-to-date and immensely informed inferences regarding behaviour and human activity that enhance the tourism industry (Xu & Whitmarsh, 2020). Through the tourist, a plethora of data sources can inform decisions made at different phases, such as before, during and after a trip. One example of a popular data source is social media, which is used by

**CONTACT** E. Rahmadian e.rahmadian@rug.nl

tourists and also tourism businesses to communicate, find or deliver relevant information or any recommendation, and to obtain information about important news and crises (Park et al., 2019).

Among the branches of tourism, sustainable tourism has understandably gained traction in recent years (Demunter, 2017). Sustainable tourism is not only about achieving economic purposes, but also assessing environmental, and social dimensions with the focus on mitigating or remediating environmental, social and economic impacts (Loaiza et al., 2019). This type of tourism also has an important role in the achievement of the 2030 Agenda of Sustainable Development Goals (SDGs) (UNWTO & UNDP, 2017). In particular, it can accelerate the accomplishment of all goals in the agenda because it has a multiplier effect on other sectors. The contribution of sustainable tourism to obtaining the goals can be direct or in-direct, specifically when it comes to Goal 8 (on inclusive and sustainable economic growth), Goal 12 (on sustainable consumption and production – SCP) and Goal 14 (on sustainable use of oceans and marine resources) (UNWTO, 2005). Considering the large playing field of sustainable tourism, big data could be an alternative input to measure sustainability. Despite these benefits, research on the use of big data for sustainable tourism is still limited.

To start filling the aforementioned gap, the work presented in this paper focuses on analysing the use of big data for sustainable tourism research by conducting a systematic literature review (SLR). In particular, we expect to show the current situation, challenges and opportunities of using big data for researchers and professionals from both the private and public sectors in the field of sustainable tourism. To that end, this study is aimed at presenting a complete analysis to understand and describe big data analytics focused on sustainable tourism by investigating four facets: (a) sources of big data, (b) approaches, (c) purposes and (d) contexts of application.

The remainder of this paper is organized as follows. In Section 2, we present the relevant background information to our SLR. In Section 3, we describe our study design, including the objectives, research questions and SLR protocol. In Section 4, we then present the results based on the primary studies that have been examined. Finally, we present the conclusions, describe the remaining work and address the threats to the validity of our study in Section 5.

## 2. Theoretical framework

### *2.1. Sustainable tourism*

The concept of sustainable tourism has been discussed for decades by concerned people aware of environmental issues such as the destruction of biodiversity, the earth-heat escalation, the diminishing of the ozone layer and the greenhouse effect; and also social issues such as discrimination among large groups of population (Richard & Derek, 2000). The United Nations Environment Programme (UNEP) and the World Tourism Organization (UNWTO) explained that while fostering tourism activities, sustainable tourism accounts for the host communities' social, cultural and economic needs, along with attempting to promote revisitations (UNWTO, 2005). The key purpose of sustainability is therefore the ability to encourage revisitation while simultaneously encouraging preservation of the social, cultural and economic characteristics of the destination and to protect the environment.

The collaboration for sustainable tourism development is an optimistic and critical task. Innovative ideas for the sustainable development of tourism are needed to learn from past mistakes, to fully seize its possibilities, to meet its responsibilities and to embrace complexity and chaos in collaborative efforts for resilient action. The latter suggests that tourism has obligations not only to itself as an industry, but also to its clients, stakeholders and staff, to governance, to people, to other nations and over time (Libburd & Edwards, 2018). Tourism is a view through which we can begin to understand modern society, e.g. in the organization of leisure time, in representations of self to others, in interpretations and in perceptions of risk and safety. Tourism figures parts of the world into sites of work and plays places to be respected and well-maintained or turned into production and development. Sustainability requires a strong focus on the tourism industry attracting visitors to the

destination free from being at the cost of the environment or society, including the local communities' attitudes (Guilarte & Gonzales, 2018).

### 2.2. Big data

In the era of digitalization and industry 4.0, data have become necessary for all aspects of life including the tourism sector. Government and all related stakeholders need data for decision-making, policy framework, business evaluation and prediction. Considering the world dynamic and technology enhancements, it is now urgent to find new data sources, e.g. by exploring big data. In the publication 'A World That Counts' (IEAG, 2014), the United Nations promoted that data play a useful role in the realization of sustainable development. In 2017, UNWTO also encouraged big data exploration for tourism statistics, even though big data can be challenging and stakeholders need to be vigilant in terms of accuracy, validity and reliability (UNWTO & UNDP, 2017).

Currently, technological developments related to big data further allow transformation and rapid innovation in tourism (Sigala, 2018). Technological tools grant the real-time, fast, and mobile capture and sharing of a big amount of data in several formats. For example, social media networks facilitate the fast virility of big data, fostering their enrichment, augmentation and change. Technologies also enable the rapid processing, visualization and analysis of big data helping and promoting decision-making in daily actions and also for defining strategies. In general, big data has commanded the creation of new technologies, methods, data capture applications, visualization methods and data aggregation abilities (Gandomi & Haider, 2015). In this line, big data is commonly designated in terms of V's: Volume, Variety, Velocity, Validity, Veracity, Value, Visibility, Visualization and Virility in spreading (Gunther et al., 2017).

In contrast to traditional data, economic development requires rich, multidimensional (i.e. economic, socio-cultural and environmental data from many stakeholders/perspectives), real-time and spatially specific data. Big data is spatial and time-based, as well as complex, i.e. generated from various users and sources. The characteristics of big data can significantly address the needs for economic development measurement (Gandomi & Haider, 2015; Lehrer et al., 2018). Analytically, technology advances (e.g. sensors, social media and web-based tools) generate data in high volumes (large-scale data), at high velocity (high-speed real-time data), in wide variety (data variability in the form of, e.g. soft and hard data, text-based data and numerical data) and with a high level of veracity (multiple interpretations and a lot of 'noise', e.g. big data quality and reliability).

## 3. Methodology

The goal of our study, as described using the goal-question-metrics approach (Basili et al., 1994), is to 'analyze existing scientific literature for the purpose of characterizing the use of big data with respect to approaches (e.g. processes, methods and tools) from the point of view of researchers and practitioners in the context of sustainable tourism'. To achieve this goal, we derived the following research questions:

RQ1: What sources of big data have been used in solutions for sustainable tourism?

RQ2: What big data approaches have been applied to sustainable tourism?

RQ3: To what purposes have the approaches been applied?

RQ4: In what contexts have the approaches been deployed?

To assess the use of big data on sustainable tourism research, we conducted an SLR using a combination of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method (Moher et al., 2015) and the guidelines proposed by Kitchenham and Charters (2007) for SLRs in Software Engineering. We elaborate our search and analysis strategies in the following subsections.

### 3.1. Sources selection and search scope

We selected six online digital libraries as search sources, namely Scopus, Wiley Online Library, Springer Link, Science Direct, IEEE Xplore and ACM Digital Library. As also discussed by Dieste et al. (2009), we selected these sources based on their timelines (publications are regularly updated), availability (full text), quality (accuracy of the search results) and versatility (export mechanism).

The search scope was limited to papers written in English and published before January 2020 (i.e. time of the initial data collection). To build the search string, we identified keywords related to the main overarching topics of our study, namely 'big data' and 'sustainable tourism'. The keywords were selected based on term similarity (e.g. 'ecotourism' for 'sustainable tourism'), as well as extracted from a quasi-gold standard, created based on the approach by H. Zhang and Babar (2010).

To create the quasi-gold standard, we searched the aforementioned digital libraries for journals that contained the terms 'tourism' or 'big data' to identify prominent venues in these fields. Based on that, we identified 13 journals and then searched them for papers that contained both 'tourism' and 'big data' in their full text. Finally, we manually inspected these studies, resulting in the selection of 48 primary studies.

The final search string, which was also validated by experts, is as follows:

('sustainable tourism' OR 'over tourism' OR 'ecotourism' OR
'responsible tourism' OR 'geo tourism'
OR 'smart tourism' OR 'e-tourism')
AND
('big data' OR 'artificial intelligence' OR 'digital' OR
'machine learning' OR 'internet of things')

### 3.2. Study selection and analysis procedure

In total, we retrieved 1448 publications from the six digital libraries. The scope and quality criteria used for the study selection are listed in Table 1. We followed the PRISMA procedure (see Figure 1) and filtered the publications as follows:

(1) in the 'Identification' step, we identified and excluded 428 duplicates;
(2) in the 'Screening' step, we excluded 766 publications based on the title and abstract;
(3) in the 'Eligibility' step, we excluded 108 publications based on the full text.

That resulted in a final set of 144 publications: 100 journal articles, 32 conference papers, and 12 book chapters. A total of 133 papers were available for download and analysis, whereas we had no access to 11 papers. The first two authors were responsible for filtering the studies independently,

**Table 1.** PRISMA inclusion criteria.

| Topic | Criterion |
|---|---|
| Field | Sustainable tourism and big data (or similar terminology), as core unit of analysis, must be mentioned in title, abstract or key words of a publication |
| Location | Studies from any geographical locations |
| Study | Only publications from peer-review academic journals, academic |
| Design | Books, conference proceeding and book chapters are accepted |
| Topic | Publications need to provide empirical evidence to support the research question (with preference for articles providing evidence to explain the use of big data to support sustainable tourism) |
| Language | Written in English |
| Date | Published up to 2019 |

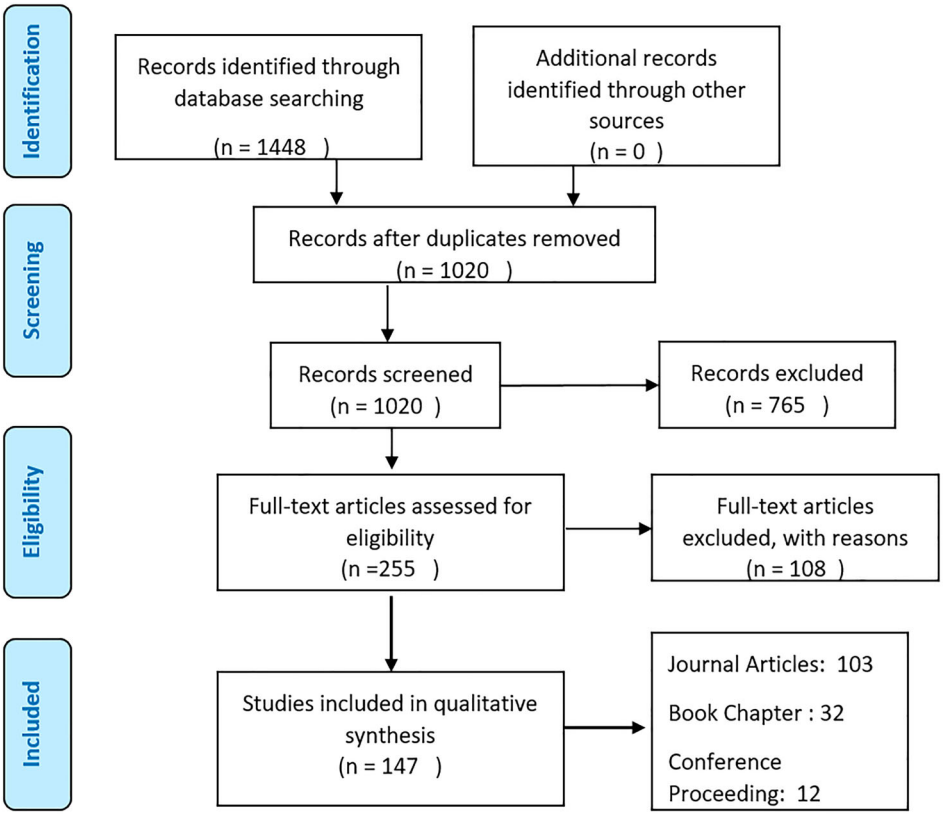


**Figure 1.** PRISMA flow diagram.

where their agreement level was calculated using Cohen's kappa, and the third author was involved in discussions regarding disagreements.

To answer the research questions, we want to examine four facets, namely (a) sources of big data, (b) approaches, (c) purposes and (d) contexts of application. Thus, from each primary result, we extracted the set of variables listed in Table 2. The variables used to answer each research question are mapped in Table 3.

## 4. Results and discussion

In this section, we present the findings based on the data analysis for each research question. First, we present the demographics of the studies included in the SLR. Then we elaborate on each facet: sources of big data, approaches, purpose and context.

**Table 2.** Extracted variables.

| Var. | Description | Var. | Description |
|---|---|---|---|
| V1 | Title | V8 | Approach's name |
| V2 | Author | V9 | Approach's source of big data |
| V3 | Year | V10 | Description of big data source |
| V4 | Type of Paper | V11 | Approach's purpose |
| V5 | Paper goal (use/propose approach) | V12 | Approach's maturity |
| V6 | Approach type (software/design) | V13 | Approach's context of application |
| V7 | Approach's sub-type | V14 | V14 Study setting (academic/non-academic) |

**Table 3.** Mapping variables to RQs.

| Research Questions | Variables |
|---|---|
| RQ1 (sources of big data) | V9, V10 |
| RQ2 (approaches) | V5–V8 |
| RQ3 (purposes) | V11, V12 |
| RQ4 (contexts) | V13, V14 |

### *4.1. Studies demographics*

Figure 2 illustrates the annual numbers of published articles on sustainable tourism research using big data. Based on the chat, there are two main conclusions that are easily derived. First, research on the use of big data on sustainable tourism is still in an early stage. On the other hand, there appears to be a generally growing trend in the annual numbers of published articles during 2016–2019, indicating increasingly wide attention to such emerging research.

Based on the type of research, most papers (58%) conducted quantitative research (e.g. statistical analysis, mathematical modelling, spatial analysis and data mining). Moreover, 34 papers (25%) conducted qualitative research (e.g. questionnaires, multifunctional analysis and descriptive analysis) and only 7 papers (5%) conducted mixed research (quantitative and qualitative). The remaining 15 papers (12%) presented theoretical research or were literature reviews.

### *4.2. Sources of big data*

This section will answer RQ1 regarding big data sources that have been used in solutions for sustainable tourism. In this section, we present and discuss the sources of big data mentioned in the collected studies. For that, we group those that are more closely related. Table 4 presents a description of the groups and the amount of studies in each group.

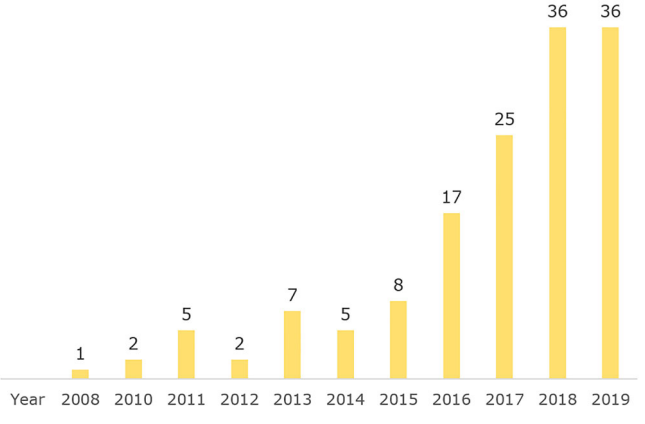


**Figure 2.** Articles by year.

**Table 4.** Number of studies per type of source.

| Source type | Description | Number of studies |
|---|---|---|
| User-generated content | Data that are actively provided by user/explicit data | 74 |
| Passive user data | Data that are collected from user's 5 information | 5 |
| Non-user data | Data that are not generated by user | 54 |

#### 4.2.1. User-generated content

Website development and then social network systems (SNSs) have facilitated tourists to share their experiences with digital world (Cong et al., 2014). In the modern tourism era, social media is inevitable for sharing experiences, photos and opinions (van der Zee & Bertocchi, 2018). As a platform, social media allows tourists to receive information and also to share their thoughts and recommendations, either positive or negative, about accommodation, places, attractions or services (Budeanu, 2013). Review websites such as Tripadvisor, Booking.com, TravelPod, Tripblog, Lyping, Weibo and Sina that promote attractions, restaurants, accommodation or used other travel services and facilities have allowed for a wider community and researchers to analyse the insights regarding comments, reviewer's profile and characteristic, trends and also how tourists' expectations are met with real experiences (Cong et al., 2014; Liu et al., 2017; Periañez-Cristobal et al., 2019; Radjevic et al., 2015; van der Zee & Bertocchi, 2018; Wang, 2018).

Tourists' digital footprints from SNSs have become valuable data sources not only for the tourism business and marketing (Onder, 2017), but also to support sustainable development through the analysis of tourists' opinion and impression of the environment, cultures, traditions and communities (Loaiza et al., 2019). Moreover, social media applications related to sustainable tourism such as Virtual Communities and Virtual Tours could support preserving the destinations that are threatened by massive and negative impacts of tourism (Budeanu, 2013; Cheng et al., 2017); to analyse tourists' behaviour in the planning stage, in the pre-trip stage, during the trip and after the trip (Hussain et al., 2018); or to understand the differences of tourists' behaviour between regions or countries (Liu, Huang, et al., 2019).

In addition, posts with embedded geo-locations enable researchers to identify visitors flow (Cvelbar et al., 2018; Nilashi et al., 2019; Tang & Li, 2016). Geo-tagged photos enabled researchers to analyse tourists' spatial network related to cultural heritages (Nguyen, Camacho, et al., 2017; Nguyen, Hwang, et al., 2017), cultural ecosystem services (CES) (Retka et al., 2019) and conservation areas (Kim, Kim, et al., 2019). The source of geo-tagged images are posted by tourists using social network platforms such as Facebook, Instagram, Twitter, Flickr, Panoramio and Foursquare (Encalada et al., 2017; Giglio et al., 2019a; Hausmann et al., 2018; Nguyen, Camacho, et al., 2017; Salas-Olmedo et al., 2018).

Further, sets of photos using GPS coordinate (latitude, longitude) were also clustered as geo-clusters to identify the densest among them (i.e. Place of Interest). Understanding geo-tagged photos at cultural heritage sites would be useful to provide tourists some useful information related to their trip (Nguyen, Camacho, et al., 2017), and enable researchers to capture the movement trajectories of tourists on a larger scale (Vu et al., 2015). Other studies have also proposed to use the city and country of origin of tourists, and discover the number of foreign and domestic tourist in the area (Giglio et al., 2019b).

#### 4.2.2. Passive user data

The destinations in which cards have been used can serve to analyse tourists' preferences on activities, e.g. archeological, art, historical, castles, science, cultural, environment, food and wine or nature (Scuderi & Nogare, 2018). Moreover, the card transactions data (on-site payments and cash withdrawals) are also utilized as an innovative data source to analyse domestic tourism dynamics anonymously as it used digital data sources which is less employed compare to traditional survey for statistics (Palop et al., 2019).

Roaming data of foreign mobile phone calls can be utilized as data source for empirical analysis of geographical, temporal and compositional dimensions on the actual visits of foreign visitors, through accessing the anonymized passive mobile positioning data of the major national mobile network operator(Bratucu & Cismaru, 2015). As a use case in Estonia has shown, it is possible to use such data source while complying with data protection and privacy requirements to reveal visitation patterns of the selected destinations on varying geographical levels (Raun et al., 2016).

Potentially, the call information can be enriched using sensors equipped in the devices, e.g. GPS, RFID, smart lock and Bluetooth (Tripathy et al., 2018).

#### 4.2.3. Non-user data

Satellite photos were used to investigate the transformational conditions of the ecosystem environments or levels of land cover within areas of recreation, conservation by years and areas as it provides temporal and spatial information (Aminu et al., 2013; Tussyadiah & Miller, 2019). The geographical information system (GIS) technology can provide qualitative and quantitative information (Aminu et al., 2013) (e.g. land use, road network and public facilities), being important to achieve optimal sustainable tourism in protected areas (Chao et al., 2011). GIS can also support decision-making to regulate the use of existing resources, tourism carrying capacity (TCC) and limits of acceptable change (LAC) (Pavon & Piña, 2018).

Another source-related SNSs that have been explored are online tourism advertisements, e.g. available on YouTube, Vimeo and Daily motion. They can point to cultural contexts in different countries that have been scrutinized through a qualitative approach based on the use of visual methods to identify sustainable tourism dimensions and discourses (Batat & Prentovic, 2014).

### 4.3. Approaches

This section will answer RQ2 regarding big data approaches that have been applied to sustainable tourism. In this section, we distinguish between methods (e.g. types of analysis) and tools that implement some of these methods. Then, once again we group the discussion based on associated topics. Table 5 presents a description of the groups and the amount of studies in each group. We clarify that the analysis of RQ2 encompasses primary studies that presented approaches (112 in total). Thus, theoretical, literature review or too qualitative papers (21 in total) were not considered.

#### 4.3.1. Methods based on statistical analysis

The use of big data in sustainable tourism research should enable local communities to use statistical modelling and forecasting to look forward to what their needs are and mitigate the negative impacts in the future (Caringal et al., 2017). Tables 6 and 7 present several statistical and spatial analysis along with their purposes regarding sustainable tourism research.

#### 4.3.2. Methods based on data mining and artificial intelligence

Aside from statistical modelling, predictive analysis and spatial analysis, artificial intelligence (AI) was also used for planning solutions on sustainable tourism. Table 8 presents several methods analysis and its purposes based on data mining and AI.

#### 4.3.3. Tools

Geo2Tag LBS (location-based services) platform is assigned under MIT open source license and some believe that it is the most popular Open Source LBS platform (Balandina et al., 2015). This platform is also used for e-tourism services and has helped tourists to see a detailed plan of a museum with marked objects and read stories that explain relations between various objects of the museum and other objects from other museums and hospitality industry spots (Balandina et al., 2015). Also focusing on user experience, ARTour, a game based on augmented reality, was used to promote

**Table 5.** Number of studies per type of source and approach.

| Source type | Method based on statistical analysis | Methods based on AI and ML | Tools and application |
|---|---|---|---|
| User-generated content | 43 | 26 | 1 |
| Passive user data | 2 | 2 | 1 |
| Non-user data | 15 | 6 | 16 |

**Table 6.** Statistical analysis and its purposes.

| Authors | Statistical methods | Purposes | Data source |
|---|---|---|---|
| Sarkar et al. (2013) | Factorial analysis of variance (ANOVA) | To reveal significant difference among image subjects and determine the visibility index of wildlife species in the regions | Wildlife photography imagery |
| Marchiori and Cantoni (2015) | Multiple linear regression | To identify factors affecting the perceived confirmation/disconfirmation of prior beliefs about a destination | UGC |
| Encalada et al. (2017) | Multiple linear regression | To identify tourists' behaviour regarding spatial association | Geo-tagged images |
| Pavon and Piña (2018) | Multi-criteria decision analysis, SWOT analysis and principal component analysis | To support strategic planning tools in accordance with priorities and available options to design better policies | Satellite imagery, conventional map |
| Lee et al. (2017) | Structural equation model | To develop a model that explains the effect of smart tourism technology attributes and destination experiences to tourist happiness | Online questionnaire on smart tourism |
| Han et al. (2018) | Structural equation model | To test the validity and reliability of measurements on using UGC for analysis | Online questionnaire on UGC |
| Bokelmann and Lessmann (2019) | Autoregressive integrated moving average | To evaluate Google Trend data as an useful predictor for tourist visit | Google Trend |
| Liu, Zhang, et al. (2019), Sun et al. (2019) | Granger causality test | To capture the relationship between search volume index (SVI) and tourist arrivals | Geo-tagged images, Google and Baidu[a] |
| Hopken et al. (2019) | Granger causality test | To forecast tourist's arrival | Google Trend |
| Liu, Zhang, et al. (2019) | Vector autoregression model (VAR) | To reveal the spatio-temporal characteristics of the tourist's visit | UGC |
| Gunter and Onder (2016); Onder (2017) | VAR in a fusion with Bayesian estimation | To forecast tourism demand | Google Analytics, UGC |
| Li and Pan (2017) | Generalized dynamic factor model | To predict tourist volume using a more comprehensive index | Geo-tagged photos |

[a]Add citation: Sun et al 2019

**Table 7.** Spatial analysis and its purposes.

| Authors | Statistical methods | Purposes | Data source |
|---|---|---|---|
| Pavon and Piña (2018) | Integration of statistical modelling and geographic information system (GIS) | To identify potential area for tourism and to avoid conflict between tourism activities and conservation | Satellite images, conventional maps, questionnaire |
| Retka et al. (2019) | Integration of statistical modelling and GIS | To analyse the distribution, qualities, characteristics and values of locations in cultural ecosystem services | Geo-tagged photos |
| Kim, Kim, et al. (2019) | Integration of statistical modelling and GIS | To quantify the proxy 'photo-user day' to estimate number of visitors | Geo-tagged images |
| Nguyen, Camacho, et al. (2017) | Haversine formula | To analyse path and fields when developing GIS application for clustering purpose | Geolocation data set |
| Chao et al. (2011) | Integration GIS and agent-based simulation | To explore phenomena of change in land use, land coverage and number of tourists | Satellite images |
| Aminu et al. (2013) | Multiple criteria evaluation (MCE) applying analytic hierarchy process | To support spatial decision for sustainable tourism planning by evaluating possible alternatives from different perspectives | Satellite images, conventional maps |
| Aminu et al. (2014) | MCE applying analytic natural process | To determine nature conservation and tourism development priorities among highland areas | Satellite images, conventional maps |

agritourism by encouraging tourists to maintain responsible environmental behaviour through active outdoor learning experiences (Garzon et al., 2018).

Internet of Things (IoT) can be very important in smart tourism to support tourist activities before, during and after the trip. Smart tourism systems can provide tour information via mobile apps based on tourism big data analytics (Kim, Kang, et al., 2019). Jeju, a special self-governing province in Korea,

**Table 8.** Methods and purposes based on data mining and artificial intelligence.

| Authors | Methods | Purposes |
|---|---|---|
| Rossetti et al. (2015) | Knowledge engineering, information extraction and machine learning (ML) | To propose a new model on processing textual reviews based on several application scenarios |
| Gunter and Onder (2016) | ML and natural language processing techniques | To analyse and visualize tourists' behaviour, origin and perspective |
| Giglio et al. (2019a) | Deep data examination through ML algorithms including image recognition, image classification and cluster analysis (unsupervised learning technique) | To determine the dynamics of tourism by identifying landmarks and finding patterns or grouping within data |
| Zhang et al. (2019) | Photo content analysis conducting computerized deep learning models | To understand tourist's perspective and behaviour and to explore tourists' cognition in the tourism destination |
| Li et al. (2017) | Adaptive neuro-fuzzy inference systems (ANFIS): a prediction ML technique | To predict the overall travellers' preferences on each cluster of accommodation |
| Nilashi et al. (2019) | A hybrid method of ML using clustering self organization map, higher order singular value decomposition and the ANFIS techniques | To handle multidimensional data in tourism contexts and to visualize the data which ease the data characteristics |
| Retka et al. (2019) | Content analysis based on text mining | To examine an array of dynamics such as reactions on social media |
| Huertas and Marine-Roig (2015, 2016), Cheng et al. (2017), Periañez-Cristobal et al. (2019) | Content analysis | To analyse what contents are published by tourist destinations or ecotourism sites through their official social media accounts and to find relational patterns within destinations. |
| Liu, Huang, et al. (2019), Chao and Lai (2015) | Content analysis | To analyse tourists' behaviour, experience, evaluations and satisfaction |
| Park et al. (2019) | Content analysis | To analyse campaigns using certain hashtag or terms |
| Tang and Li (2016) | Social network analysis | To analyse the relationship network and evaluate the network structure through a series of indicators |
| Onder (2017), Loaiza et al. (2019) | Semantic-based natural language data mining | To compare tourists' expectations and real experiences and to assess contrasts between tourists' expectations and service providers' based on SNS content |

deploys free WiFi access points on major tourist spots and buses to provide free internet access service via one-time authentication. This platform provides data to analyse tourists' routes, which are very important for policy establishment and personalized smart tourism services (Kim, Kang, et al., 2019). In addition, these smart tourism monitoring devices also measure the number of visitors and environmental data including dust amount, temperature, wind speed and snowfall (Kim, Kang, et al., 2019).

iTour is a GPS-enabled device equipped with Internet and other services, including RFID, smart lock, Bluetooth, smart map and sensors (Tripathy et al., 2018). It was used to improve convenience and flexibility in the tourism domain and integrated various tourist services (Tripathy et al., 2018).

CASTA (Contention, Convenience, and Accessibility based Smart Tourism-destination Approach) is a smart system which is started by the tourist handheld GPS-enabled device and allows the tourist to send a preferences message and get back the required feedback. It is also equipped each tourism site with small wireless sensors placed in various sections, rooms, squares that collect various information that could be of interest to tourists such as the temperature, humidity, air quality and visitors' number (Almobaideen et al., 2016). Similarly, sensor network on the road for enhanced internet of touristic things (STREET) Web is a platform used to distribute tourist information 'on the road' by using Smart Box (SB), a physical architecture of micro servers placed in touristic areas with monuments. This application enables interaction between visitors and objects where tourists also could access any information related to the objects (Angelaccio & Buttarazzi, 2016).

Accordingly, an IoT-based system for the smart museum was implemented to provide the user with a real interactive experience specifically for personalized navigation. By utilizing BLE (Bluetooth low energy) and computer vision techniques, tourists can obtain information related to the artworks (Fiore et al., 2016). Finally, by applying AI techniques to an Integrated Urban Tourism called Puglia Project, an e-tourism application so-called Semantic Web was developed. Semantic Web aims to provide information related to alternative transportation to the objects compliant to tourists' profiles. Therefore, tourist acts not only as information users but also information suppliers (Lisi & Esposito, 2015).

### 4.4. Purposes and context of application

This section will answer RQ3 and RQ4 regarding purposes and contexts that have been deployed towards sustainable tourism. The results suggest that most studies (approx. 29%) report studies on tourist behaviour and flow, followed by smart tourism and ICT (approx. 24%), managing sustainable tourism in general (approx. 14%), destination image/attractiveness (approx. 10%), nature preservation/conservation (approx. 9%) and ecotourism (approx. 3%). Not only discussing one specific topic or context, approximately 10% papers discussed two topics in one paper. These papers, for instance, combination of ecotourism and destination image, ecotourism and smart tourism, destination image and managing sustainable tourism in general or destination image and tourist flow.

#### 4.4.1. Ecotourism, nature preservation, cultural heritage

Ecotourism is an aspect of tourism that supports socio-economic, socio-cultural and socio-political development without decreasing the quality of the environment (Rauf et al., 2015) in which ecotourism grasps the principles of sustainable tourism (Hernandez-Aguilar et al., 2017). Analysis of climate conditions, seasonality, security and political instability has a significant influence on the visitors' arrivals/patronage of the ecotourism industry, for example at Saiful Malook National Park in Pakistan (Rauf et al., 2015). Understanding a protected area's visitation status is critical, as it strongly affects the sustainability of natural resources and environment protection (Kim, Kang, et al., 2019). Even though spatial visitation patterns and their features on nature-based tourism are difficult to assess using only a field-based survey (Kim, Kim, et al., 2019). Social big data has been used to investigate nature-based tourism by identifying locations and the reasons tourists visit to evaluate which features are attractive to tourists (Kim, Kim, et al., 2019); investigating which activities people engage with when visiting protected areas (Hausmann et al., 2018); investigating the use of social media by ecotourism management agencies and how this potentially changes the relationship between the ecotourism and the natural environment (Cheng et al., 2017). This has also been done in the Chengdu Research Base of Giant Panda Breeding (CRBGP) in Sichuan, China by examining wildlife tourism experiences in a unique site combined with a species: giant pandas (Cong et al., 2014).

Sustainable use of tourism resources should be ensured by appropriate development and utilization considering environment protection and nature conservation, which can be delivered by installing a smart system to measure the number of visitors, and environmental data of sightseeing spots such as fine dust amount, temperature and humidity, wind speed and snow fall which utilized a big data analysis in Jeju, Korea (Kim, Kang, et al., 2019). To identify conservation and compatible areas for tourism development especially in nature conservation and wetland areas, spatial modelling in GIS has been delivered in Johor Ramsar sites, Malaysia (Aminu et al., 2013). Spatial analysis was also applied to regulate the tourist and recreational activities within the Lagunas de Montebello National Park in Spain, to assess TCC and LAC to design better policies and strategies linked to tourism and sustainability (Pavon & Piña, 2018).

CES are determined as the non-material advantages derived from human–ecosystem relationships (Retka et al., 2019). To quantify the benefits of CES in spatial large area in Brazil's largest marine-protected area, which is difficult to convey, an assessment using user-contributed geo-referenced photographs from a popular image and video-hosting website has been implemented (Retka

et al., 2019). Similarly, in Vietnam, using geo-tagged resources from social networking services which combined semantic tags and media data such as image data, an analysis has been employed to bring out useful information for the purposes of smart cultural tourism development (Nguyen, Hwang, et al., 2017). On the other hand, from the cultural heritage tourism industry side, a multiple case study was applied on two leading platforms involved in the online dissemination of cultural heritage: Europeana and Google Arts & Culture, as a representative for digitization and datification of artefacts and cultural heritage (Pesce et al., 2019).

#### *4.4.2. Destination image/attractions*

Destination managers are devoting time and money to promote their destination became popular among tourist and internet users (Inversini et al., 2010), using the strategic system, for instance, consumer-oriented intelligent support system in the form of web service incorporated by tourism domains such as sightseeing spot suggestion (Chao & Lai, 2015). Understanding target-market tourist is also important to adjust the destination management objects (DMO) marketing strategies based on tourist's perspectives about destination (Deng et al., 2019; Mariani et al., 2016). Tourists' origin, nationalities, domestic vs. international, will influence their perceptions, choices and behaviour about destinations (Hopken et al., 2019; Liu, Huang, et al., 2019). Table 9 presents a more complete picture about the use of big data analysis for DMO, competitiveness and tourist destination.

#### *4.4.3. Tourist behaviour*

Understanding tourists' intrinsic interest in high-quality tourism environments, hence inspiring them to respond with a tremendous sense of personal responsibility toward the environment, could be critical to promote sustainable tourism (Han et al., 2018). Visions of tourist travel behaviours are important for managers engaged in sustainable tourism industry, to provide practical implications for destination development, transportation planning and impact management (Vu et al., 2015). This insight could be obtained using social media to analyse, four stages of travelling: planning stage, pre-trip stage, during the trip stage and post-trip stage (Hussain et al., 2018), as tourists generate a large amount of 'digital footprint' that can be used for studying tourist behaviour (Tang & Li, 2016). Traveller service satisfaction and travel experience satisfaction also has a significant effect on tourist happiness (Lee et al., 2017). Thus, the user's feedback can be used for other online users to determine and make decision to find the most relevant facilities, for instance, eco-friendly hotels, tailored to their preferences (Nilashi et al., 2019). Ability to analyse travel behaviours will also enable the

**Table 9.** Big data Analysis for DMO, competitiveness and tourist destination.

| Authors | Purposes | Methodology | Data Sources |
|---|---|---|---|
| Inversini et al. (2010) | To promote destination among tourist and internet users | Content analysis | UGC |
| Chao and Lai (2015) | To promote destination among tourist and internet users | Consumer-oriented intelligent support system | Web service incorporated with tourism domain |
| Marchiori and Cantoni (2015) | To deliver information of products and services to tourists and specific audience | Sentiment analysis | UGC |
| Deng et al. (2019) | To adjust DMO marketing strategies based on tourist's origin and perspective on destination | Content analysis | Geo-tagged UGC |
| Hopken et al. (2019) | To understand tourist's origin, nationalities, domestic vs. international which influence perspectives and choices | Regression analysis | UGC |
| Liu, Huang, et al. (2019) | To understand different characteristics of domestic tourist compared to international visitors | Sentiment analysis | UGC |
| Raun et al. (2016) | To identify which places are popular among different nationalities in Estonia | Statistical and spatial analysis | Mobile positioning data |
| Qi and Chen (2019) | To assess Macao's image between Chinese tourists and international tourists as it implies segmentation and marketing strategies | Content analysis | UGC |

provider to adopt different strategies in designing and promoting their tourism products (Michopoulou & Moisa, 2016).

Visitor flows in tourism research could be employed as proxies for tourism demand (Cvelbar et al., 2018). Hence, flows are measured as number of tourist's arrivals and receipts from different countries to the host destination (Cvelbar et al., 2018). However, the use of big data to estimate or predict tourism demand is a new technique and also an important step for tourism industry and research (Cvelbar et al., 2018), where capability to forecast tourism demand accurately will ease destination management organization to provide enough facilities, information and transportation system (Onder, 2017). Moreover, the travel route would be utilized as the basic for tourism cooperation, tourism planning, tourism marketing, travel route design and other purposes (Tang & Li, 2016). Ability to predict tourism demand will implicate to efficient allocation, cost reduction and to detect tourist behaviour in the early stage (Gunter & Onder, 2016).

However, there is little evidence on the spatial behaviour of urban tourists, even though tourists generate large data when they visit cities (Salas-Olmedo et al., 2018). The analysis of user-generated content (UGC) in Antwerp and the example of the MAS museum highlight difficulties associated with tourist flow over space in historic cities (van der Zee & Bertocchi, 2018). Meanwhile, visual content analysis of tourist photos is considered to be an effective way to analyse tourist's cognition in the tourism destination (Zhang et al., 2019). Not only UGC and photos, destination cards and the use of credit cards are also beneficial to analyse and identify the most sequence of tourist activities, which implied to the policy of marketing strategies of the destination (Scuderi & Nogare, 2018).

#### *4.4.4. Smart tourism*

Tourism, as one of the fastest-growing industries, has played an important role in the context of Smart Cities and sustainability, as tourism is a considerably important generator of carbon emissions (Nitti et al., 2017). Smart tourism is therefore an important part of smart cities (Tripathy et al., 2018). Recently, smart tourism has emerged as a new term to illustrate the advanced technology applications that rely on sensors, big data processing technique, open data, open Application Programming Interface (API), new way of connecting and exchanging between humans and machines and multi-device (such as IoT, Radio-frequency Identification (RFID) and Near-field Communication (NFC)) in tourism (Vu et al., 2018). Smart tourism also relies on the ability to not only collect enormous amounts of data, but to intelligently store, process, combine, analyse and use big data to inform business innovation, operations and services by AI and big data techniques (Tsaih & Hsu, 2018). There are several implementations of smart tourism in the context of sustainable tourism, first, the use of AI techniques to support integrated tourism services in Apulia Region, Italy (Lisi & Esposito, 2015). These integrated services consists of several tools such as OnTourism, with data automatically retrieved from the websites of Trip Advisor and Google Maps, which lead into actions oriented to service innovation for the sustainable knowledge society (Lisi & Esposito, 2015).

Second, a design of a no-invasive indoor location-aware architecture, which improves the user experience in a museum, in Lecce (Italy) (Fiore et al., 2016). An IoT-based system for smart tourism services has been implemented towards smart museum navigation which improves visitor's cultural experience (Fiore et al., 2016).

Third, a system (BRB (Be Right Beach)) advocates avoiding the overcrowding of beaches by allowing each tourist to choose the right beach to visit to have the best experience (Girau et al., 2018). This system consists of control units equipped with an UV sensor, thermometers, a humidity sensor and a camera to predict the crowd. Hence, this system provides information and suggestion for tourists based on tourists' preferences and complete information (Girau et al., 2018).

Fourth, utilizing the smart tourism ecosystem model as the framework, using SNS social computing formula as the theoretical basis, it has been established the Living Green tourism ecosystem, which enable to enhance relation between business and the smart travel itinerary recommendation in order to support improvement of the entire industry (Tsai et al., 2018).

Further, the smart cultural tourism service would convey smart interactions between the tourists themselves by collecting and analysing geo-tagged datasets from available social media (Nguyen, Camacho, et al., 2017). Also, to support mobile users of smart tourism, for instance, smart museum, STREET Web has been utilized by integrating micro servers, distributed in the scenario (servers on the road), called SBs and functioning as a geo-based Cloud system in an autonomous way, as a Distributed Local Storage system, without remote internet access (Angelaccio & Buttarazzi, 2016).

#### *4.4.5. Managing sustainable tourism in general*

Unsustainable tourism development will lead to problems such as loss of natural resources, conflicts between tourists and local residents and so on (Chao et al., 2011). As a part for the policy-making process to achieve sustainable tourism, especially in conservation areas, big data analysis, AI, GIS and the IoT are utilized (Chao et al., 2011). For instance, adopting social media could be employed for sustainable tourism even though slowly adopted by institutional actors (Budeanu, 2013). A conceptual framework for sustainability defined as systematic thinking for sustainable tourism, which can solve complex problems by applying intra- and intertextual analysis, has been applied by using online tourism advertisement as data sources (Batat & Prentovic, 2014). This technique might help stakeholders to understand cultural differences while promoting sustainable tourism in certain areas (Batat & Prentovic, 2014).

Another opportunity to enhance the quality of sustainable tourism management is by using social media which enables individuals to deliver their thoughts to tourism decision-making and planning, even though its adoption is relatively slow by institutional actors (Budeanu, 2013). Not only for the policy-making process, social media such as Facebook has also been utilized to engage students as a medium of education in the context of sustainable tourism, where students not only collected information, but also actively created and shared the knowledge about sustainability (Isacsson & Gretzel, 2011). Social media also plays a role in crisis communication strategy in the tourism sector as it is open to the public, interactive and traceable (Park et al., 2019). Social media is deeply changing the way tourists search, find, use and collaboratively produce information about tourism, in which this participatory approach will enable the merge of traveller's needs and the local communities in sustainable development (Guzzo et al., 2013), and also to link the tourists, industry and environmental responsibility (Kleinrichert et al., 2012).

## 5. Related works and limitations

There are several papers discussing the use of social media for sustainable tourism, for instance, identifying the most important issues related to Social Media Analytics and Smart Tourism (SMAST) and offer some guidelines for future research through an SLR (Viñan-Ludeña, 2019). A range of the academic reports were also reviewed in order to derive the crucial information and to have a better overview of ICT-based tools/applications in all the dimensions of ecotourism (Katsoni & Dologlou, 2017), analysis of smart technology that exploits IoT for the travel and hospitality industry (Wise & Heidari, 2019), analysis of tourism big data for forecasting (Song & Liu, 2017), literature review on the use of big data for tourism research (J. Li et al., 2018), and analysis of smart tourism development as a major focus for many destinations across the world particularly in Asia (Gretzel et al., 2018). Our paper discussed a wide range analysis of methods, tools and approach towards sustainable tourism to give broader perspective on how big data and AI are utilized over sustainability.

Regarding the threats to the validity of our study, we considered the following aspects. First, the search string may not be optimal and miss relevant keywords and, therefore, papers. To mitigate this threat, we created a quasi-gold standard to ensure the quality of the search string and relevance of the selected papers. Second, to mitigate risks related to data collection and analysis, in addition to the details provided in Section 3, there were thorough discussions among the three authors along with all steps of the SLR so far. Third, we did not employ the snowballing technique to trace other

possibly related papers. We also acknowledge that, given inherent subjective biases, replications of this research could lead to slightly different select studies and results.

## 6. Conclusion

This paper reports the study design and findings of an SLR to identify the existing and potential uses of big data for sustainable tourism applications. Big data research on sustainable tourism seems to be emerging since 2016. In an attempt to start answering our research questions, the results showed that the use of big data for sustainable tourism is research in both academic (106 papers) and non-academic settings (25 papers), and that approaches serve for various tasks, mostly for identification, ranking and prediction of behaviours as well as number of tourists (Garau, 2017; Kenterelidou et al., 2017; Kim & Kim, 2017). In addition, big data is used to support decision and policy-making processes (Loureiro, 2018; Zhang, F, 2019), as well as for services provision based on a variety of data sources (Zhang, 2016). To support the decision-making, it is important to include its integration with technologies (e.g. AI-based approaches and IoT) in the application designing (Kim, Kang, et al., 2019) to support smart tourism as one of important elements in smart cities and sustainability (Fiore et al., 2016; Girau et al., 2018; Lisi & Esposito, 2015; Nguyen, Camacho, et al., 2017). Table 5 suggests that UGC and non-user data were deployed in various methods such as statistical analysis, AI and ML, also as data sources for some tools and applications.

Although the use of big data in research has advanced, it still poses challenges concerning data quality, data cost and privacy concerns (Li et al., 2018). Regarding the quality, big data could cover a large sample, but not necessarily the whole sample (Xu & Whitmarsh, 2020). There are also several issues related to the reliability of UGC and online review data, for instance, fake reviews for negative purposes (Xu & Whitmarsh, 2020). Further, in a view of privacy concerns, mobile roaming and transaction data have not been widely used in tourism research yet are expensive and difficult to access, despite the advantages. It occurred in the matter of fact that tourists, mobile operator networks, also business transaction providers do not wish to share their private information (Raun et al., 2016; Palop et al., 2019).

To overcome those challenges, mutual cooperation between academia, government and industries becomes essential. This cooperation will not only ensure data availability and reduce data cost for tourism research using big data, but also addresses practical issues in return. Furthermore, having confidentiality agreements and excluding sensitive information might be another solution to solve the problem of privacy concerns.

Regarding quality challenges, we also aim to collaborate on future opportunities with practitioners and other stakeholders, especially to take advantage of various research methods, data sources, approaches and designs. Finally, we aspire to investigate to what extent the primary studies address paramount implications to end-users such as privacy and security, and to utilize big data analysis in cyber-physical system towards sustainability.

## Acknowledgements

Authors extend their gratitude to the editor and the reviewers for their valuable comments and insights on the manuscript.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

*E. Rahmadian* http://orcid.org/0000-0002-5052-7703
*D. Feitosa* http://orcid.org/0000-0001-9371-232X
*A. Zwitter* http://orcid.org/0000-0001-7391-141X

## References


Almobaideen, W., Allan, M., & Saadeh, M. (2016). Smart archaeological tourism: Contention, convenience and accessibility in the context of cloud-centric IoT. *Mediterranean Archaeology and Archaeometry*, *16*(1), 227–236. https://doi.org/10.5281/zenodo.35535

Aminu, M., Ludin, A. N. B. M., Matori, A. N., Yusof, K. W., Dano, L. U., & Chandio, I. A. (2013). A spatial decision support system (SDSS) for sustainable tourism planning in Johor Ramsar sites, Malaysia. *Environmental Earth Sciences*, *70*(3), 1113–1124. https://doi.org/10.1007/s12665-012-2198-6

Aminu, M., Matori, A., & Yusof, K. (2014). A spatial decision support system (SDSS) for sustainable tourism planning in Cameron Highlands, Malaysia. *IOP Conference Series: Earth and Environmental Science*, *18*(1), 1–8. https://doi.org/10.1088/1755-1315/18/1/012139

Angelaccio, M., & Buttarazzi, B. (2016, April 23-25). *Street web: A framework for "web on the road" smart services*. Webist 2016 - Proceedings of the 12th International Conference on Web Information Systems and Technologies (WEBIST), Rome, Italy.

Balandina, E., Balandin, S., Koucheryavv, Y., & Mouromtsev, D. (2015, November 23-27). *Innovative e-tourism services on top of Geo2Tag LBS platform*. 11th International Conference on Signal-Image Technology and Internet-based Systems (SITIS 2015), Bangkok, Thailand (pp. 752–759).

Basili, V. R., Caldiera, G., & Rombach, H. D. (1994). Goal question metric paradigm. In J. J. Marciniak (Ed.), *Encyclopedia of software engineering*. Wiley & Sons (pp. 528–532).

Batat, W., & Prentovic, S. (2014). Towards viral systems thinking: A cross-cultural study of sustainable tourism ads. *Kybernets*, *43*(3/4), 529–546. https://doi.org/10.1108/K-07-2013-0147

Bokelmann, B., & Lessmann, S. (2019). Spurious patterns in Google trends data - an analysis of the effects on tourism demand forecasting in Germany. *Tourism Management*, *75*, 1–12. https://doi.org/10.1016/j.tourman.2019.04.015

Bratucu, G., & Cismaru, L. (2015). Developing a business intelligence planning tool for managing ecotourism destinations based on indicators existing at EU level. *International multidisciplinary scientific GeoConference Surveying Geology and Mining Ecology Manage ment (SGEM)*, *3*(5), 181–188.

Budeanu, A. (2013). Sustainability and tourism social media. In *Tourism social science series* (pp. 87–103). Emerald Group.

Caringal, A. M., Boongaling, J. A., Arcega, J. D., & Gonzales, R. R. (2017). From logging to blogging: Ecotourism in the mountains of lobo, Verde island passage, Batangas province, Philippines. *Asia Life Science*, *26*(2), 257–282.

Chao, A. F. Y., & Lai, C.-Y. (2015). *SNS opinion-based recommendation for etourism: A Taipei restaurant example* (Vol. 51, pp. 59–70) Springer.

Chao, D., Furuta, K., & Kanno, T. (2011). Agent-based simulation system for supporting sustainable tourism planning. In T. Onada, D. Bekki, & E. McCready (Eds.), *Lecture notes in Computer science (including subseries Lecture notes in Artificial intelligence and Lecture notes in Bioinformatics)* (Vol. 6797, pp. 243–252). Springer.

Cheng, M., Wing, I. A., & McDonald, M. (2017). Ecotourism social media initiatives in China. *Journal of Sustainable Tourism*, *25*(3), 416–432. https://doi.org/10.1080/09669582.2016.1214141

Cong, L., Wu, B., Morrison, A. M., Shu, H., & Wang, M. (2014). Analysis of wildlife tourism experiences with endangered species: An exploratory study of encounters with giant pandas in Chengdu, China. Tourism Management, *40*, 300–310. https://doi.org/10.1016/j.tourman.2013.07.005

Cvelbar, L. K., Mayr, M., & Vavpotic, D. (2018). Geographical mapping of visitor flow in tourism: A user-generated content approach. *Tourism Economics*, *24*(6), 701–719. https://doi.org/10.1177/1354816618776749

Demunter, C. (2017, June 21-23). *Tourism statistics: Early adopters of big data?* 6th UNWTO International Conference on Tourism Statistics (pp. 1–2), Manila, Phillipines.

Deng, N., Liu, J., Dai, Y., & Li, H. (2019). Different cultures, different photos: A comparison of Shanghai's pictorial destination image between east and west. *Tourism Management Perspectives*, *30*, 182–192. https://doi.org/10.1016/j.tmp.2019.02.016

Dieste, O., Grim'an, A., & Juristo, N. (2009). Developing search strategies for detecting relevant experiments. *Empirical Software Engineering*, *14*(5), 513–539. https://doi.org/10.1007/s10664-008-9091-7

Encalada, L., Boavida-Portugas, I., Ferreira, C. C., & Rocha, J. (2017). Identifying tourist places of interest based on digital imprints: Towards a sustainable smart city. *Sustainability (Switzerland)*, *9*(12). https://doi.org/10.3390/su9122317

Fiore, G. D., Mainetti, L., Mighali, V., Patrono, L., Alleto, S., Cucchiara, R., & Serra, G. (2016). A location-aware architecture for an IoT-based smart museum. *International Journal of Electronic Government Research*, *12*(2), 39–55. https://doi.org/10.4018/IJEGR.2016040103

Gandomi, A., & Haider, M. (2015). Beyond the hype big data concepts methods and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Garau, C. (2017). Emerging technologies and cultural tourism: Opportunities for a cultural urban tourism research agenda. In N. Bellini & C. Pasquinelli (Eds.), *Tourism in the city* (pp. 67–80). Springer International Publishing.

Garzon, J., Acevedo, J., Pavon, J., & Baldiris, S. (2018). ARTour: Augmented reality-based game to promote agritourism. In L. T. De Paolis & P. Bourdot (Eds.), *Augmented reality, virtual reality, and computer graphics* (pp. 413–422). Springer International Publishing.

Giglio, S., Bertacchini, F., Bilotta, E., & Pantano, P. (2019a). Machine learning and point of interests: Typical tourist Italian cities. *Current Issues in Tourism*, *23*(13), 1646–1658. https://doi.org/10.1080/13683500.2019.1637827

Giglio, S., Bertacchini, F., Bilotta, E., & Pantano, P. (2019b). Using social media to identify tourism attractiveness in six Italian cities. *Tourism Management*, *72*, 306–312. https://doi.org/10.1016/j.tourman.2018.12.007

Girau, R., Ferrara, E., Pintor, M., & Sole, M. (2018, 30 July-3 Aug). *Be right beach: A social IoT system for sustainable tourism based on beach overcrowding avoidance*. Proceedings - IEEE 2018 International Congress on Cybermatics: 2018 IEEE Conferences on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, iThings/GreenCom/CPSCom/SmartData/BlockChain/CIT 2018 (pp. 9–14).

Gretzel, U., Ham, J., & Koo, C. (2018). Creating the city destination of the future: The case of smart Seoul. In Y. Wang, A. Shakeela, A. Kwek, & C. Khoo-Lattimore (Eds.), *Managing Asian destination* (pp. 199–214). Springer.

Guilarte, P., & Gonzales, L. (2018). Sustainability and visitor management in tourist historic cities: The case of Santiago de Compostela, Spain. *Journal of Heritage Tourism*, *13*(6), 489–505. https://doi.org/10.1080/1743873X.2018.1435665

Gunter, U., & Onder, I. (2016). Forecasting city arrivals with Google analytics. *Annals of Tourism Research*, *61*, 199–212. https://doi.org/10.1016/j.annals.2016.10.007

Gunther, W., Mehrizi, M. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information*, *26*(3), 191–209. https://doi.org/10.1016/j.jsis.2017.07.003

Guzzo, T., D'Andrea, A., Ferri, F., & Grifoni, P. (2013). A framework to promote and develop a sustainable tourism by using social media. *Lecture Notes in Computer Science*, *8186*, 656–665. https://doi.org/10.1007/978-3-642-41033-8_83

Han, W., McCabe, S., Wang, Y., & Chong, A. Y. L. (2018). Evaluating user-generated content in social media: An effective approach to encourage greater pro-environmental behavior in tourism? *Journal of Sustainable Tourism*, *28*(4), 500–614. https://doi.org/10.1080/09669582.2017.1372442

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., & di Minin, E. (2018). Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conservation Letters*, *11*(1), Article e12343. https://doi.org/10.1111/conl.12343

Hernandez-Aguilar, J. A., Cortina-Villar, H. S., García-Barrios, L. E., & Castillo-Santiago, M. A. (2017). Factors limiting formation of community forestry enterprises in the southern Mixteca region of Oaxaca, Mexico. *Environmental Management*, *59*(3), 490–504. https://doi.org/10.1007/s00267-017-0821-8

Hopken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2019). Google trends data for analysing tourists' online search behaviour and improving demand forecasting: The case of ARE, Sweden. *Information Technology and Tourism*, *21*(1), 45–62. https://doi.org/10.1007/s40558-018-0129-4

Huertas, A., & Marine-Roig, E. (2015). Destination brand communication through the social media: What contents trigger most reactions of users?. In I. Tusyadiah & A. Inversini (Eds.), *Information and communication technologies in tourism 2015* (pp. 295–308). Springer International Publishing.

Huertas, A., & Marine-Roig, E. (2016). Differential destination content communication strategies through multiple social media. In A. Inversini & R. Schegg (Eds.), *Information and communication technologies in tourism* (pp. 239–252). Springer International Publishing.

Hussain, T., Li, B., & Wang, D. (2018). What factors influence the sustainable tour process in social media usage? Examining a rural mountain region in Pakistan. *Sustainability (Switzerland)*, *10*(7). https://doi.org/10.3390/su10072220

IEAG. (2014). *A world that counts: Mobilising the data revolution for sustainable development*. UN Data Revolution.

Inversini, A., Marchiori, E., Dedekind, C., & Cantoni, L. (2010). Applying a conceptual framework to analyze online reputation of tourism destinations. In U. Gretzel, R. Law, & M. Fuchs (Eds.), *Analysis and modeling of complex data in behavioral and social sciences* (pp. 321–332).

Isacsson, A., & Gretzel, U. (2011). Facebook as an edutainment medium to engage students in sustainability and tourism. *Journal of Hospitality and Tourism Technology*, *2*(1), 81–90. https://doi.org/10.1108/17579881111112430

Katsoni, V., & Dologlou, N. (2017). ICT applications in smart ecotourism environments. In A. Stratigea, E. Kyriakides, & C. Nicolaides (Eds.), *Smart cities in the Mediterranean* (pp. 563–575). Springer International Publishing.

Kenterelidou, C., Galatsopoulou, F., & Skamnakis, A. (2017). Agrotourism and well-being sustainability: A communication approach to 'what matters and better life'. *Rivista di Studi sulla Sostenibilita*, *1*(1), 129–147. https://doi.org/10.3280/RISS2017-001009

Kim, D., & Kim, S. (2017). The role of mobile technology in tourism: Patents, articles, news, and mobile tour app reviews. *Sustainability (Switzerland)*, *9*(11). https://doi.org/10.3390/su9112082

Kim, J. S., Kang, H. J., & Jwa, J. W. (2019). Smart tourism monitoring device based on IoT technologies. *International Journal of Recent Technology and Engineering*, *7*(6), 1281–1285.

Kim, Y., Kim, C. K., Lee, D. K., Lee, H. W., & Andrada, R. I. T. (2019). Quantifying nature-based tourism in protected areas in developing countries by using social big data. *Tourism Management*, *72*, 249–256. https://doi.org/10.1016/j.tourman.2018.12.005

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (EBSE Technical Report).

Kleinrichert, D., Ergul, M., Johnson, C., & Uydaci, M. (2012). Boutique hotels: Technology, social media and green practices. *Journal of Hospitality and Tourism Technology*, *3*(3), 211–225. https://doi.org/10.1108/17579881211264495

Lee, J., Lee, H., Chung, N., & Koo, C. (2017). An integrative model of the pursuit of happiness and the role of smart tourism technology: A case of international tourists in Seoul. In R. Schegg & B. Stangl (Eds.), *Information and communication technologies in tourism 2017* (pp. 173–186). Springer International Publishing.

Lehrer, C., Brocke, J., Jung, R., & Seidel, S. (2018). How big data analytics enables services innovation: Materiality, affordance, and the individualization of service. *Journal of Management Information Systems*, *35*(2), 424–460. https://doi.org/10.1080/07421222.2018.1451953

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*(6), 301–323. https://doi.org/10.1016/j.tourman.2018.03.009

Li, N., Tung, V., & Law, R. (2017). A fuzzy comprehensive evaluation algorithm for analyzing electronic word-of-mouth. *Asia Pacific Journal of Tourism Research*, *22*(6), 592–603. https://doi.org/10.1080/10941665.2017.1308395

Li, X., & Pan, B. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, *59*, 57–66. https://doi.org/10.1016/j.tourman.2016.07.005

Libburd, J., & Edwards, D. (2018). *Collaboration for sustainable development*. Goodfellow Publishers Ltd.

Lisi, F. A., & Esposito, F. (2015). An ai application to integrated tourism planning. In M. Gavanelli, E. Lamma, & F. Riguzzi (Eds.), *AI*IA 2015 advances in artificial intelligence* (Vol. 9336, pp. 246–259). Springer International Publishing.

Liu, P., Zhang, H., Zhang, J., Sun, Y., & Qiu, M. (2019). Spatial-temporal response patterns of tourist flow under impulse pre-trip information search: From online to arrival. *Tourism Management*, *73*, 105–114. https://doi.org/10.1016/j.tourman.2019.01.021

Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management*, *71*, 337–347. https://doi.org/10.1016/j.tourman.2018.10.004

Liu, Z., Shan, J., Glassey, B., & Fang, G. (2017). Semantic social media analysis of Chinese tourists in Switzerland. *Information Technology and Tourism*, *17*(2), 183–202. https://doi.org/10.1007/s40558-016-0066-z

Loaiza, J. P. V., Pérez-Torres, A., & Contreras, K. M. D. (2019). Semantic icons: A sentiment analysis as a contribution to sustainable tourism. *Sustainability (Switzerland)*, *11*(17). https://doi.org/10.3390/su11174655

Loureiro, A. (2018). There is a fourth industrial revolution: The digital revolution. *Worldwide Hospitality and Tourism Themes*, *10*(6), 740–744. https://doi.org/10.1108/WHATT-07-2018-0044

Marchiori, E., & Cantoni, L. (2015). The role of prior experience in the perception of a tourism destination in user-generated content. *Journal of Destination Marketing and Management*, *4*(3), 194–201. https://doi.org/10.1016/j.jdmm.2015.06.001

Mariani, M. M., Felice, M. D., & Mura, M. (2016). Facebook as a destination marketing tool: Evidence from Italian regional destination management organizations. *Tourism Management*, *54*, 321–343. https://doi.org/10.1016/j.tourman.2015.12.008

Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. John Murray.

Michopoulou, E. E., & Moisa, D. (2016). Information and communication technologies in tourism. *Information and Communication Technologies in Tourism*, 765–777. https://doi.org/10.1007/978-3-319-28231-2_55

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Stewart, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews J*, *4*(1), 1–9. https://doi.org/10.1186/2046-4053-4-1

Nguyen, T. T., Camacho, D., & Jung, J. E. (2017). Identifying and ranking cultural heritage resources on geotagged social media for smart cultural tourism services. *Personal and Ubiquitous Computing*, *21*(2), 267–279. https://doi.org/10.1007/s00779-016-0992-y

Nguyen, T. T., Hwang, D., & Jung, J. J. (2017). Using geotagged resources on social media for cultural tourism: A case study on cultural heritage tourism. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (LNICST)*, *194*, 64–72. https://doi.org/10.1007/978-3-319-58967-1_8

Nilashi, M., Ahani, A., Esfahani, M. D., Yadegaridehkordi, E., Samad, S., Ibrahim, O., Sharef, N. M., & Akbari, E. (2019). Reference learning for eco-friendly hotels recommendation: A multi- criteria collaborative filtering approach. *Journal of Cleaner Production*, *215*, 767–783. https://doi.org/10.1016/j.jclepro.2019.01.012

Nitti, M., Pilloni, V., Giusto, D., & Popescu, V. (2017). IoT architecture for a sustainable tourism application in a smart city environment. *Mobile Information Systems*. https://doi.org/10.1155/2017/9201640

Onder, I. (2017). Forecasting tourism demand with Google trends: Accuracy comparison of countries versus cities. *International Journal of Tourism Research*, *19*(6), 648–660. https://doi.org/10.1002/jtr.2137

Palop, J. D. R., Arias, J. M., Bodas-Sagi, D. J., & Lapaz, H. V. (2019). Determining the usual environment of cardholders as a key factor to measure the evolution of domestic tourisms. *Information Technology and Tourism*, *21*(1), 23–43. https://doi.org/10.1007/s40558-018-0130-y

Park, D., Kim, W. G., & Choi, S. (2019). Application of social media analytics in tourism crisis communication. *Current Issues in Tourism*, *22*(15), 1818–1824. https://doi.org/10.1080/13683500.2018.1504900

Pavon, R. G. S., & Piña, I. I. C. (2018). Sustainable tourism development in Parque Nacional Lagunas de Montebello, Mexico. *WIT Transactions on Ecology and the Environment*, *227*, 83–93. https://doi.org/10.2495/ST180091

Periañez-Cristobal, R., Luengo, P., & Calvo-Mora, A. (2019). The heritage character and service quality in urban defensive structures renovated as hotels. *Journal of Heritage Tourism*, *15*(1), 44–59. https://doi.org/10.1080/1743873X.2019.1590375

Pesce, D., Neirotti, P., & Paolucci, E. (2019). When culture meets digital platforms: Value creation and stakeholders' alignment in big data use. *Current Issues in Tourism*, *22*(15), 1883–1903. https://doi.org/10.1080/13683500.2019.1591354
Qi, S., & Chen, N. C. (2019). Understanding Macao's destination image through user-generated content. *Journal of China Tourism Research*, *15*(4), 503–519. https://doi.org/10.1080/19388160.2019.1577199
Radjevic, T., Stanisic, N., & Stanic, N. (2015). Solo travellers assign higher ratings than families: Examining customer satisfaction by demographic group. *Tourism Management Perspectives*, *16*, 247–258. https://doi.org/10.1016/j.tmp.2015.08.004
Rauf, T., Yukan, C., Zada, M., Khan, N., & Shah, S. J. (2015). Impact of Saiful Malook national park on the sustainable livelihood of Naran and Kaghan communities, Pakistan. *GeoJournal*, *85*, 1227–1239. https://doi.org/10.1007/s10708-019-10019-z
Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, *57*, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006
Retka, J., Jepson, P., Ladle, R. J., Malhado, A. C. M., Vieira, F. A. S., Normande, I. C., Correia, R. A., Souza, C. N., Bragagnolo, C., & Correia, R. A. (2019). Assessing cultural ecosystem services of a large marine protected area through social media photographs. *Ocean and Coastal Management*, *176*, 40–48. https://doi.org/10.1016/j.ocecoaman.2019.04.018
Richard, D., & Derek, H. (2000). *Tourism and sustainable community development*. Routledge.
Rossetti, M., Stella, F., Cao, L., & Zanker, M. (2015). Analysing user reviews in tourism with topic models. In I. Tusyadiah & A. Inversini (Eds.), *Information and communication technologies in tourism 2015* (pp. 47–58). Springer International Publishing.
Salas-Olmedo, M. H., Moya-Gomez, B., Garćıa-Palomares, J. C., & Gutíerrez, J. (2018). Tourists' digital footprint in cities: Comparing big data sources. *Tourism Management*, *66*, 13–25. https://doi.org/10.1016/j.tourman.2017.11.001
Sarkar, S. K., Au, N., & Law, R. (2013). Analysing ecotourists satisfaction in socialisation and knowledge sharing intentions via social media. In Z. Xiang & I. Tussyadiah (Eds.), *Information and communication technologies in tourism 2014* (pp. 313–326). Springer International Publishing.
Scuderi, R., & Nogare, C. D. (2018). Mapping tourist consumption behaviour from destination card data: What do sequences of activities reveal? *International Journal of Tourism Research*, *20*(5), 554–565. https://doi.org/10.1002/jtr.2205
Sigala, M. (2018). New technologies in tourism: From multi-disciplinary to anti-disciplinary advances and trajectories. *Tourism Management Perspective*, *25*, 151–155. https://doi.org/10.1016/j.tmp.2017.12.003
Song, H., & Liu, H. (2017). Predicting tourist demand using big data. In Z. Xiang & D. R. Fesenmaier (Eds.), *Analytics in smart tourism design* (pp. 13–29). Springer International Publishing.
Sun, S., Wei, Y., Tsui, K.-L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, *70*, 1–10. https://doi.org/10.1016/j.tourman.2018.07.010
Tang, J., & Li, J. (2016). Spatial network of urban tourist flow in Xi'an based on microblog big data. *Journal of China Tourism Research*, *12*(1), 5–23. https://doi.org/10.1080/19388160.2016.1165780
Tripathy, A. K., Tripathy, P. K., Ray, N. K., & Mohanty, S. P. (2018). iTour: The future of smart tourism: An IoT framework for the independent mobility of tourists in smart cities. *IEEE Consumer Electronics Magazine*, *7*(3), 32–37. https://doi.org/10.1109/MCE.2018.2797758
Tsai, T.-H., Chang, T.-H., Lin, Y.-W., Yu, M.-C., Lien, P.-J., Yan, W.-C., & Ho, W.-L. (2018). Emerging social media and social networks analysis transforms the tourism industry: Living green smart tourism ecosystem. In M. Antona & C. Stephanidis (Eds.), *Universal access in human-computer interaction. Virtual, augmented, and intelligent environments*. Springer International Publishing (Vol. 10908, pp. 583–590).
Tsaih, R.-H., & Hsu, C. C. (2018, December 2-6). *Artificial intelligence in smart tourism: A conceptual framework*. International Conference on Electronic Business (ICEB), Guangxi, China (pp. 124–133).
Tussyadiah, I., & Miller, G. (2019). Perceived impacts of artificial intelligence and responses to positive behaviour change intervention. In J. Pesonen & J. Neidhart (Eds.), *Information and communication technologies in tourism 2019*. Springer International Publishing (pp. 359–370).
UNWTO. (2005). Making tourism more sustainable: a guideline for policy makers. World Tourism Organization Pubns.
UNWTO, & UNDP. (2017). *Tourism and the sustainable development goals-journey to 2030*. World Tourism Organization Pubns.
van der Zee, E., & Bertocchi, D. (2018). Finding patterns in urban tourist behaviour: A social network analysis approach based on TripAdvisor reviews. *Information Technology and Tourist*, *20*(1-4), 153–180. https://doi.org/10.1007/s40558-018-0128-5
Viñan-Ludeña, M.-S. (2019). Systematic literature review on social media analytics and smart tourism. In V. Katsoni & M. Segarra-Oña (Eds.), *Smart tourism as a driver for culture and sustainability* (pp. 357–374). Springer International Publishing.
Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management*, *46*, 222–232. https://doi.org/10.1016/j.tourman.2014.07.003
Vu, T. H., Dai, N. H. B., ThiKhank, N. H., & Quang, D. N. V. (2018, December 6-7). *Overall structural system solution for supporting services and tourists management oriented on smart city in Vietnam*. ACM International Conference Proceeding Series, Da Nang, Vietnam (pp. 430–435).

Wang, P., Zhang, X., Suomi, R., & Sun, C. (2018). Determinants of customers' eWOM behaviour—a system success perspective. In R. Schegg & B. Stangl (Eds.), *Information and communication technologies in tourism 2017* (pp. 401–415). Springer International Publishing.

Wise, N., & Heidari, H. (2019). Developing smart tourism destinations with the internet of things. In *Big data and innovation in tourism, travel, and hospitality: Managerial approaches, techniques, and applications* (pp. 21–29). Springer Nature.

Xu, F., & Whitmarsh, L. (2020). Big data or small data? A methodological review of sustainable tourism. *Journal of Sustainable Tourism*, *28*(2), 144–163. https://doi.org/10.1080/09669582.2019.1631318

Zhang, F. (2019, May 10-12). *Prediction and evaluation of urban eco-sports tourism behavior using data mining technology*. ACM International Conference Proceeding Series (Vol. 19, pp. 595–608).

Zhang, H., & Babar, M. A. (2010, April 12-13). *On searching relevant studies in software engineering*. 14th International Conference Evaluation and Assessment in Software Engineering, Keele University, Keele, UK (pp. 111–120).

Zhang, K., Chen, Y., & Li, C. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tourism Management*, *75*, 595–608. https://doi.org/10.1016/j.tourman.2019.07.002

Zhang, Q. (2016). Probability analysis of tourists revisit intention based on data mininbg techniques. *International Journal of Simulation: Systems, Science and Technology*, (17), 4.1–4.6. https://doi.org/10.5013/IJSSST.a.17.17.04