

Big data in tourism research: A literature review

Jingjing Li ^a, Lizhi Xu ^{b, c}, Ling Tang ^{a, *}, Shouyang Wang ^{d, e}, Ling Li ^a

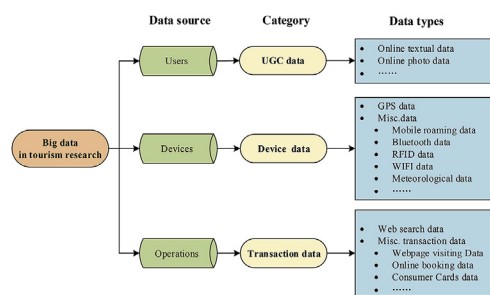


^a School of Economics and Management, Beihang University, Beijing 100191, China
^b Collaborative Innovation Center of eTourism, Beijing Union University, Beijing 100101, China
^c Tourism College, Beijing Union University, Beijing 100101, China
^d Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
^e School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China

HIGHLIGHTS

- Various big data have been applied to tourism research, making a great improvement.
- A comprehensive review on different types of big data in tourism is presented.
- Carrying different information, different types address different tourism issues.
- Each type's research focuses, data characteristics and analytic tools are analyzed.
- The corresponding major challenges and further directions are further investigated.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
 Received 13 September 2017
 Received in revised form 24 January 2018
 Accepted 8 March 2018

Keywords:
 Tourism research
 Big data
 Literature review
 Tourism management
 Tourist behavior

ABSTRACT

Even at an early stage, diverse big data have been applied to tourism research and made an amazing improvement. This paper might be the first attempt to present a comprehensive literature review on different types of big data in tourism research. By data sources, the tourism-related big data fall into three primary categories: UGC data (generated by users), including online textual data and online photo data; device data (by devices), including GPS data, mobile roaming data, Bluetooth data, etc.; transaction data (by operations), including web search data, webpage visiting data, online booking data, etc. Carrying different information, different data types address different tourism issues. For each type, a systematical analysis is conducted from the perspectives of research focuses, data characteristics, analytic techniques, major challenges and further directions. This survey facilitates a thorough understanding of this sunrise research and offers valuable insights into its future prospects.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

With the rapid development of computer science and Internet techniques, massive-scale data in both structured and unstructured styles are generated, recorded, stored and accumulated, forming the *big data* and opening a new age (Kambatla, Kollias, Kumar, & Grama, 2014). In such a big data era, a variety of big data,

* Corresponding author. School of Economics and Management, Beihang University, 37 Xueyuan Road, Haidian District, Beijing 100191, China.
 E-mail address: lingtang@buaa.edu.cn (L. Tang).

together with the conceptual and technological innovations, have been used in extensive areas of science, engineering, healthcare, management, business, tourism, etc. (Hashem et al., 2015). However, there has not existed a uniform definition of big data yet, with different researchers giving a variety of definitions. A famous and original definition is the 3V, which characterizes big data as *Volume*, *Variety* and *Velocity* (Laney, 2001). Gantz and Reinsel (2011) extended the 3V's definition to a 4V's concept, by introducing *Value* to highlight the validity and usefulness of big data. Even though controversial in definition, big data and big data technologies have made great contributions to improving scientific researches, with tourism research as an emerging, typical example.

Even at an early stage, a rich mine of tourism-related big data have been generated from three primary sources—users, devices and operations. First, the Internet has fostered a rapid rise in social media, offering a capacious platform to spread user-generated content (UGC) data (in terms of texts, photos, etc.) (Xiang, Du, Ma, & Fan, 2017). Second, due to the vigorous development of Internet of things (IoT), diverse sensor devices have been developed and employed to track tourist movements and environmental conditions, providing considerable spatial-temporal big data (such as global position system (GPS) data, mobile roaming data, Bluetooth data, etc.) (Shoval & Ahas, 2016). Third, tourism is a complex system covering a series of operations (i.e., transactions, activities or events in tourism market) such as web searching, webpage visiting, online booking & purchasing, etc., thus producing the corresponding transaction data of web search data, webpage visiting data, online booking data etc., for understanding tourist behavior and improving tourism marketing. Based on the big data from these three main sources, tourist behavior and tourism market can be better explored and understood by both academia and industries.

Using the aforementioned valuable big data, tourism research has stepped into the big data era and brought forth amazing improvements. For instance, Yang, Pan, Evans, and Lv (2015) advocated that the large scale of big data could finely make up for the limitation of sample size issues faced by survey data users, providing a new way to understand tourist behavior. Similarly, Li, Pan, Law, and Huang (2017) argued that big data analytics could provide sufficient data without sampling bias, helping both academia and industries better understand tourist behavior. Xiang, Schwartz, Gerdes, and Uysal (2015) claimed that big data analytics could develop new knowledge to reshape the understanding of hospitality industry and to support the corresponding decision making. With the aforementioned superiorities, big data allowed a better understanding of tourism demand, tourist behavior, tourist satisfaction and other tourism issues.

Given that big data have substantially changed the traditional tourism research based on traditional data, a comprehensive review on such sunrise research, i.e., using big data in tourism research, is extremely desired. On the one hand, although facing a growing volume of publications, this sunrise research field was still not very clearly known by potential researchers. For example, it was still uncertain about what particular types of big data have been used in tourism and how to take advantage of these new data. On the other hand, when comparing big data and traditional data, the former might be much more informative and structure-complex, thus appearing different data characteristics, focusing on different research issues, and requiring different analytic techniques. Therefore, a comprehensive review on full-scale types of big data in tourism research is strongly required in terms of research focuses, data characteristics and analytic techniques, in order to present a historical tour of how these particular big data have contributed to tourism research and to offer helpful insights into the future prospects.

However, a systematic literature review on the big data in tourism research was still lacking. The existing literature reviews on tourism

research mainly focused on the following issues: tourism within a certain country such as China (e.g., Bao, Chen, & Ma, 2014; Huang & Chen, 2016; Sun, Wei, & Zhang, 2017; Zhang, Lan, Qi, & Wu, 2017); tourism in particular types such as event tourism and volunteer tourism (Getz & Page, 2016; Wearing & McGehee, 2013); tourism demand (Goh & Law, 2011; Song & Li, 2008); tourist behavior (Bhati & Pearce, 2016; Pomfret & Bramwell, 2016); tourism attraction (Leask, 2016); tourism risk (Yang, Khoo-Lattimore, & Arcodia, 2017). Nevertheless, regarding applying big data to tourism research, to the best of our knowledge, there existed only three relevant reviews: Rashidi, Abbasi, Maghrebi, Hasan, and Waller (2017) explored the capacity of social media data for modelling travel movements; Schuckert, Liu, and Law (2015b) reviewed the studies on online reviews in tourism and hospitality; Shoval and Ahas (2016) conducted a literature review on tracking data in tourism research. Obviously, the three studies focused on a certain type of big data (social media data, online reviews, or tracking data), without an overall analysis on full-type big data. Moreover, all the three studies were conducted mainly from the dimension of research fields, without a full consideration of data characteristics and analytic techniques. However, a different type of big data (from a different data source and with different information) is certainly different from any of other types, in terms of research focuses (on different tourism issues), data characteristics and processing techniques. Therefore, this paper attempts to fill in such a literature gap to present a comprehensive literature review on different types of big data in tourism research, and provide a systematical analysis on each type from the perspectives of research focuses (on tourism issues), data characteristics, analytic techniques, challenges and further directions.

The main goal of this paper is to present a comprehensive literature review on the application of big data to tourism research. Relative to the existing studies, the major contributions of this paper can be summarized into three aspects: (1) to the best of our knowledge, it might be the first attempt to review the full-scale types of big data used in tourism research; (2) given that a different type of big data (carrying different information) might address different tourism issues, appear different data characteristics and require different analytic techniques, a systematical analysis for each type is conducted from the perspectives of research focuses (on tourism issues), data characteristics and analytic techniques; (3) based on such a thorough survey, the major challenges and future prospects are carefully investigated.

The remaining part of this paper is organized as follows. Section 2 presents the general findings (or statistics) of the reviewed literature, together with the analytical framework of this paper. By following this framework, Sections 3–5 thoroughly investigate the big data in tourism research derived from the three primary sources, i.e., users, devices and operations, respectively. Section 6 concludes the main findings of the review and points out the further directions of applying big data to tourism research.

2. General findings

This section displays the general findings (statistics) of the review. First, Section 2.1 presents the literature collection. Section 2.2 provides a descriptive statistical analysis of the selected literature. Finally, Section 2.3 formulates the general analytical framework of this paper.

2.1. Databases

The articles on tourism research using big data are collected from the following academic databases: Web of Science, ScienceDirect, SAGE Journals Online, Emerald Insight, Springer and Wiley Online Library. Additionally, the powerful search engine, Google

Download English Version:

<https://daneshyari.com/en/article/7420873>

Download Persian Version:

<https://daneshyari.com/article/7420873>

[Daneshyari.com](https://daneshyari.com)