5-2017

A Systematic Literature Review: Existing Hospitality & Tourism Research on Big Data

Jiaojiao Song
University of Nevada, Las Vegas

A Systematic Literature Review:

Existing Hospitality & Tourism Research on Big Data

By

Jiaojiao Song

University of Nevada, Las Vegas

May 2017

A professional paper submitted in partial fulfillment

of the requirements for the

Master of Science in Hotel Administration

William F. Harrah College of Hotel Administration

Graduate College

University of Nevada, Las Vegas

May 2017

Chair: Dr. Mehmet Erdem

PART ONE

Introduction

Big Data has been in the eye of storm in all walks of live. Since the tourism industry is able to generate tons of data each day, Big Data has caught the attention of scholars and professionals of the industry. People might hear of Big Data here and there, yet are they clear about what Big Data means? Does Big Data refer to a certain amount data or a technology to acquire information to support evidenced decision?

Unfortunately there is no general consensus in academic area according to the book, *Big Data: Principals and Paradigms* (Buyya, Calheiros, & Dastjerdi, 2016) . The authors held the idea that the term Big Data itself was not clear. Although it did reveal the size of data, it didn't define how "big" is enough to be categorized as Big Data, which was supposed to be concerned with "time, space and circumstance". Further review on research is needed to define the term.

Different from that of decades ago, when paper took the role of primary data storage, data today are comprised of not only textual information but also visual and audio data, a large part of which come from community websites. The tourism industry is facing such a great opportunity that a large number of these community websites are related to tourism industry and could be made use of to dig out information to support strategic business decisions.

Despite such large amount of data ready to be analyzed and travelers' attention on the information provided by community websites, a large of companies in tourism industries have not taken measures to take advantage of big data.

With a systematic review on existing tourism & hospitality research on Big Data, questions listed above might be able to find potential answers.

Purpose

The purpose of this paper is to analyze and summarize existing tourism and hospitality research on Big Data through conducting a systematic literature review.

Statement of Objective

As an exploratory study, this paper is aimed at inspiring the hospitality industry on how to take advantage of the huge amount of data generated from daily operations and social networks as other industry might already have achieved. During this process, future field of study will also be excavated as existing studies will be collected and examined.

Justifications

As is mentioned, Big Data has been ubiquitous in both industry and academic world. The results of related academic studies are fruitful. While in industry, SAS, a major software company based in the United States who provides analytics, business intelligence and data management company to other companies, puts it this way: it is not the amount of data that is important; it is what the organizations do with the data that matters. Studies are needed to inspire business how to utilize the data available to support their business decisions.

Although it is undeniable that businesses like airline ticketing and hotel booking websites have been using data to perform dynamic pricing to optimize revenue, there still remains a lack of innovation with regard to the use of big data by hospitality industry. Online sectors of the industry, inclusive of online travel agencies, meta-search engines, and some information company of travel distributions, have stepped forward in this area, while other companies need to make a move in the near future if they are not willing to be knocked down by new

business models like Airbnb's peer-to-peer online marketplace and homestay network (Shafiee & Ghatari, 2016) .

Constraints

The primary constraint of current study lies in the fact that the definitions of the term Big Data are various and consequently the literature review may not include a whole body of published studies regarding the topic. Another limitation might come from the finite quantity of existing studies which restricts the width and depth of this paper.

PART TWO

Introduction

In this section, first previous studies are reviewed to define the definition of Big Data, followed by a systematic literature review conducted on 44 articles published by IEEE during 2006 to 2016 (Cacho, Estaregue, Figueredo, et al., 2015; Petroni, Biagi, Colonnese, Cusani, & G. Scarano, 2015; Sato, 2015; see Appendix A for full article list) . These publications were retrieved from IEEE Xplore Digital Library using “tourism” as the first keyword and “Big Data” as the second. Then abstracts were read to pick out those studies not directly related to hospitality and tourism research on Big Data but contain the two keywords in the articles.

Defining Big Data

Now it’s time to solve the question “what is Big Data” in academic realm based on the opinions of Gandomi and Haider (2015). They summarized the traits of Big Data into three dimensions, volume, variety and velocity.

Volume represents the magnitude of data and is impacted by time and the type of data. Given the 100% growth rate of data noticed by the aforementioned librarian, the volume of data today is 16 times of that in 2012, let alone 70 years ago when the data exploitation was first came into notice. Large data set of that time may be easily stored in a flash disc today and a common data set of today may be beyond a scientist's imagination of that time. The type of data is another factor. Even though two datasets take up exactly the same space of storage, the amount of information they convey could vary based on the data types, such as textual data v.s. video data. Hence, different techniques of storage and management and analytical methods are required to deal with distinct types of data, though they all fall into the category of Big Data.

Variety is the second dimension, referring to the structural heterogeneity in a dataset. Data could be grouped into structured data, semi-structured data as well as unstructured data. Structured data are those stored in a spreadsheet or a relational database. Such characteristic allows traditional statistical method to analyze structured data, so they were the primary type of data utilized by Big Data. Yet they only occupy 5% of data existed. On the opposite, unstructured data are not blessed with a structural organization necessary for conventional analysis, such as text, pictures, and video that cannot be placed in a table. In between the structured and unstructured data are the semi-structured data. The Extensible Markup Language, or XML in short, is a type of semi-structured data famous for encoding a document in format that can be read by both human and machine.

Velocity describes the speed of data generation and analysis, and companies' actions on them. Due to the popularization of smart phones, sensors and similar digital devices, data are generated at an unparalleled speed and bring the necessity of real-time analytics and evidence-based planning. Business runners may need real0time analytics to handle with huge amount of data, where Big Data technologies could outperform traditional data management systems in that the latter are not able to deal with high volume of data streams instantly, which might be valuable but also perishable.

In addition to these three dimensions, some other standards deserve some attention as well. IBM introduced Veracity as the fourth V to depict the unreliability of the data. Data such as customer reviews are uncertain as they involve personal judgments, but they may still contain meaningful information and thus should be paid attention to. SAS added Variability to quantify the variation of data flow rates, which usually are not constant. Last but not least,

Value was marked by Oracle.

Buyya et al. (2016) urged to consider all attributes of Big Data from different definitions, each of which might focus on either single issue or data aspects and neglect other crucial attributes. He summarized such definitions into 7 popular definitions of Big Data, namely original 3Vs definition, which has been introduced above; a technology oriented definition inclusive the development and utilization of new technology such as MapReduce and so on; Big Data as applications of various types of data, such as process-mediated data, human-sourced information and machine-generated data; Big Data as signals for new trends; Big Data as potential opportunities; Big Data as a process of human thinking; and Big Data as "a new bottle for old wine". After reviewing all popular definition of Big Data raised by previous studies, Buyya attempted to define the term from three different domains, namely Data domain to search for patterns, Business Intelligence domain for prediction making, and statistical domain for assumptions making, expanding the original 3V to 3^2V shown in Figure 1.

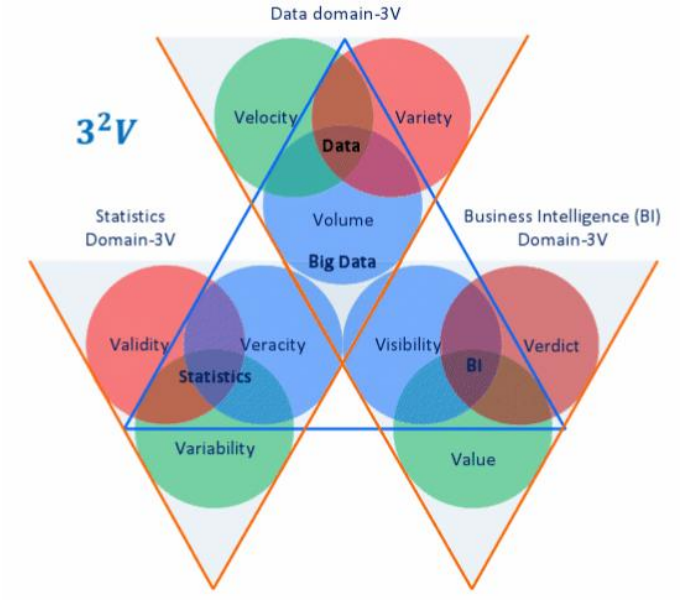

Figure *1.* 3^2 Vs Venn diagram in hierarchical model. Reprinted from *Big data: Principles and paradigms* (P.14), Buyya, R., Calheiros, R. N., & Dastjerdi, A. V., 2016, Saint Louis, Missouri: Elsevier Science.

In their opinion, the aforementioned 3Vs fell in data domain. The statistical domain used Veracity, Validity and Variability to describe the trustworthiness of the data set and reliability of the data sources. If the data themselves are not reliable, the prediction derived from them may hardly be trustworthy. In the Business Intelligent domain, Value, Visibility and Verdict were introduced to measure the motivations of implementing Big Data Analytics process at the first place. Of the three domains, the core dimensions were Volume, Veracity and Visibility of each domain, which can be seen in the center triangle in the Venn diagram.

In their article, Höpken and Fuchs (2016) implied Big Data processes involved first extracting information from various data sources, then analyzing data using techniques and analytics, and last but not least gaining new knowledge to support decision making. Gandomi and Haider (2015) held a similar view and summarized the Big Data Processes in a flow chart (Figure 2).

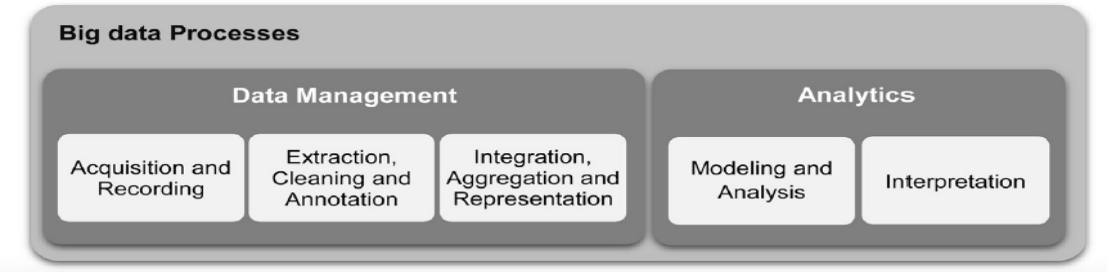

Figure 2. Flow Chart of Big Data Process. Reprinted from "Beyond the hype: Big data concepts, methods, and analytics," by Gandomi, A., & Haider, M., 2015, International *Journal of Information Management, 35*(2), 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

A Brief History of Big Data

A Forbes contributor, Press (2013) traced the origin of Big Data back to 1944, when the librarian of Wesleyan University published "The Scholar and the Future of the Research Library", first attempting to quantify the growth rate of the volume of data. According to his estimation, American university libraries had been multiplying by two in volume every 16 years. Seventeen years later, Price recorded his conclusion about the growth rate of the quantity of new journals in his publication that such growth was exponential instead of linear, which was named as "Law of Exponential Increase".

In 1967, two scholars suggested that the storage requirements for all information be kept to a minimum because of "information explosion" in a paper titled "Automatic Data Compression". Later in 1971, the trend was noticed that people were measured by the size of storage capacity of their dossiers. Four years later, Japanese Government began to track the volume of information circulating nationwide, discovering that the accretion rate of information supply surpassed that of information consumption considerably and that the demand for mass media originated data remained standstill yet that of data from personal telecommunication media had been going up dramatically. Similarly, the Hungarian Central

Statistics Office launched a study in 1981 focusing on its information industry, involving measuring information volume (as sited in Press, 2013).

In August 1983, Science published an article, “Tracking the Flow of Information”, recording the development of 17 major communication media during the period of 1960 and 1977. It also presented that the driving force of growth had transferred from broadcasting to point-to-point media by the end of observation period. In 1996, Morris and Truskowski stated that when it came to storing data, digital storage had become more cost-effective than paper. One year later, Cox and Ellsworth published the article “Application-controlled Demand Paging for Out-of-core Visualization” in the IEEE eighth conference on Visualization, being the first article to use the term “Big Data” in the Association for Computing Machinery (ACM) digital library. In the same year, Lesk published another article, “How Much Information Is There in the World?” forecasting that all information would be able to be stored as a result of the production of tape and disk, even though he estimated there might be some thousand petabytes of information by that year (as sited in Press, 2013).

In 1998, “The Size and Growth Rate of the Internet” was published by Coffman and Odlyzko, who believed that data traffic in the United States would exceed voice traffic in 2002 and be dominated by the Internet. Odlyzko also started to track the growth in Internet traffic from 2002 to 2009 through establishing the Minnesota Internet Traffic Studies. In 1999, “Visually Exploring Gigabyte Data Sets in Real Time” was published by a group of scholars, becoming the first Communications of the ACM (CACM) article to utilize the tem “Big Data”. In the article, powerful computers were regarded as a blessing to fields of inquiry as well as a curse as they had been generating massive amounts of data (as sited in Press, 2013).

Coming to the year of 2000, Lyman and Varian from UC Berkeley published the first comprehensive study to quantize the total amount of new and original information titled "How Much Information?" They also found that 92% of the new information generated in 2002, which was estimated to be 5 exabytes, was stored on magnetic media like hard disks in a study conducted in 2003 (as sited in Press, 2013).

One conclusion could be drawn from the Press's view of Big Data history that the volume of data and the capacity of data storage have been growing rapidly and new media have taken the place of traditional media in terms of major driving force of the increment of data volume.

Analytics and Applications

Data are valuable only when they are transferred into meaningful information through efficient processes and then support business decision making process. Gandomi and Haider (2015) split such process into two parts, data management and analytics. Analytics are those techniques data scientists used to analyze data and extract useful information from them. They roughly categorized Big Data Analytics into Text Analytics, Audio Analytics, Video Analytics, on the basis of data type, as well as Social Media Analytics and Predictive Analytics.

Text analytics, or text mining, are able to analyze textual data retrieved from social network feeds, emails, blogs, online forums, survey responses, corporate documents, news and even call center logs. Such analytics comprise statistical analysis, computational linguistics and machine learning, extracting meaningful summaries from high volume of text generated by human. A good example could be predicting the fluctuations of stock market

through the information provided by financial news. Similarly, tourism operators may use text analytics to dig out new insight about interest and behavior of their customers, derived on which new product and service could be designed. Audio analytics may be utilized as Interactive Voice Response (IVR) platforms to deal with callers, while video analytics may find its strength in helping retail outlets collect demographic information such as age, gender and ethnicity and purchasing behaviors of customers through analyzing video stream recorded by CCTV cameras in stores.

Social media analytics has been a focus of the industry as a result of the giant growth of social media and the popularity of UGC on the Internet, from which experiences, opinions and feelings of customers can be reached and analyzed directly (Shafiee & Ghatari, 2016) . The resource of data to be analyzed includes social networks, blogs, microblogs, social news like Reddit, media sharing sites like Instagram and YouTube, Wikis, question-and-answer sites like Quora and of course review sites like Yelp and TripAdvisor. Predictive analytics, as its name reveals, are used to predict future outcomes. Techniques like moving averages struggle to discover patterns from historical data and then apply to the future. On the other hand, techniques may also be employed to examine the interdependencies among variables, namely explained variables and explanatory variables, and utilize such interdependencies to make predictions. Business runners could make use of online analytical services, including price prediction and ranking, to maximize shopping probability (Gandomi & Haider, 2015) .

Systematic Literature Review

Literature Description

IEEE stands for the Institute of Electrical and Electronics Engineers, a professional

association promoting the engineering process of “creating, developing, integrating, sharing, and applying knowledge about electro and information technologies and sciences for the benefit of humanity and the profession”. The reason for selecting IEEE Xplorer Digital Library as the main database lies in the fact that Big Data itself is highly related to information technology and IEEE has held numerous conferences calling attention on Big Data of both scholars’ and practitioners’. The number of publications finally retrieved from the database also displays the rationality of such choice.

Conference is the main publishing channels, while some articles are published on journals as well. Table 1 shows the publication titles of the articles.

Table 1

Publication Counts By Titles

Publication Title	Count
2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications	1
2008 4th International Conference on Wireless Communications, Networking and Mobile Computing	1
2009 42nd Hawaii International Conference on System Sciences	1
2009 International Conference on New Trends in Information and Service Science	1
2010 18th International Conference on Geoinformatics	1
2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)	1
2010 Seventh International Conference on Information Technology: New Generations	1
2013 IEEE International Conference on Big Data	1
2013 International Conference on Advances in ICT for Emerging Regions (ICTer)	1
2014 IEEE International Conference on Big Data (Big Data)	1
2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS)	1
2015 IEEE First International Conference on Big Data Computing Service and Applications	3
2015 IEEE First International Smart Cities Conference (ISC2)	1
2015 IEEE International Conference on Big Data (Big Data)	2
2015 Intelligent Systems and Computer Vision (ISCV)	1

Publication Title	Count
2015 International Conference on Intelligent Transportation, Big Data and Smart City	4
2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS)	1
2016 10th International Conference on e-Commerce in Developing Countries: with focus on e-Tourism (ECDC)	1
2016 12th World Congress on Intelligent Control and Automation (WCICA)	1
2016 16th International Symposium on Communications and Information Technologies (ISCIT)	1
2016 22nd International Conference on Automation and Computing (ICAC)	1
2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)	1
2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)	1
2016 6th International Conference - Cloud System and Big Data Engineering (Confluence)	1
2016 IEEE International Conference on Big Data (Big Data)	3
2016 IEEE International Conference on Big Data Analysis (ICBDA)	1
2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)	1
2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech)	1
2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)	1
2016 International Conference on Big Data and Smart Computing (BigComp)	1
2016 International Conference on Smart Grid and Electrical Automation (ICSGEA)	1
2016 International Symposium on Computer, Consumer and Control (IS3C)	1
IEEE Access	1
IEEE Transactions on Big Data	1
International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06)	1
OCEANS 2015 - Genova	1
Total	44

Articles were published on 34 distinct conferences and journals across 2007 to 2016. Specifically, the publishing years of these articles are summarized in Table 2. No related study was found published before 2006 in IEEE Xplorer Digital Library.

Table 2 presents that an abrupt increase of number of publications on tourism research on Big Data could be seen in 2015. Previously the number of such publications each year was no more than three. Yet in the year of 2015, this number soared to 13, more than the total of the

past seven years, and later in 2016, the number reached 20, which is about half of the sum of studies published during 2007 and 2016 on IEEE conferences and journals.

Table 2

Publication Count by Year

Year	Count
2006	1
2007	1
2008	1
2009	2
2010	3
2013	2
2014	1
2015	13
2016	20
Total	44

Although the theme of a study may focus on various aspects and the introduction of a certain technology or system is always followed by its applications, here in the literature review, articles are categorized roughly into two groups based on its main purpose - whether to explain the principle of a technology and the process of utilizing it, or to emphasize its possible applications in tourism industry, as is showed in Table 3. For example, an article introducing how to construct E-Tourism platform based on service value broker (Duan, Wei, Kattepur, & Du, 2013) would fall into the technology category, even though it also involved

applications of such platform in real business world. Another example would be a case study of 2014 FIFA World Cup host city as a smart destination (Cacho, et al., 2015). This study would be labeled as “application” as its purpose was to present how other big event could apply such initiative despite the fact that the authors spared no effort in articulating the mechanism of such smart city.

Table 3

Publication Counts by Main Theme

Main Theme	Count
Application	21
Technology	23
Total	44

As IEEE is focused more on the technology side, it was anticipated that the majority of the publications would be tagged as technology oriented. Yet it turns out to be not the real case. Table 3 indicates that the new technology and the application are almost equally crucial in researchers’ opinion.

Further, if combined with the year of publishing, this will give us information about the trend of research focus, as is presented in Table 4.

Table 4

Publication Count by Main Theme & by Year

Main Them / Year	Count
Application	
2007	1
2009	2
2010	1

Main Them / Year	Count
2013	1
2014	1
2015	10
2016	6
Technology	
2006	1
2008	1
2010	2
2013	1
2015	3
2016	14
Total	44

The first eight years in question are relatively balanced between the focus on application and on technology. However, in the year of 2015, applications caught the attention of scholars, surpassing technologies. Not lasting for long, this situation is reversed in 2016, during which 14 articles were published on introducing new technologies.

Key Findings

Key findings from studies examined will be discussed in this section. These findings include both the possible applications of Big Data technologies in tourism industry and the new technologies that might be applied to tourism industry in the near future.

Among the application-oriented studies, numerous applications of Big Data technology were proposed by scholars. Making predictions to support business decisions has been mentioned frequently. In a study published in 2009, data mining and questionnaires were adopted to evaluate risk prediction of legalizing gambling in Taiwan when the bill was proposed to allow gambling in subsidiary islands of Taiwan (Shia, Li, Chen & Hsu, 2009) . In another study, data mining, data management, semantic engine and prediction in Big Data were utilized to analyze the impact of geographic environment on a province in China, where

tourism has been one of the main industries (Xiong & Wu, 2015) . Specifically, hotels could forecast occupancy and make daily sales plan using history data and the model fitted by Tsuda, Ando and Ichifuji (Tsuda, Ando, & Ichifuji, 2016) .

Big Data being used to support real-time decision-making process is another popular topic. A group of scholars proposed that the log data of WI-FI access points at tourist spots could be tracked and visualized to obtain the movement history of individual visitors to take the place of questionnaires and interviews as data resource so that both money and time could be saved (Ichifuji, Matsuo, Koide, et al., 2016). Similarly, WI-FI data could also be used to estimate the number of tourism in real time (Koide, Ichifuji, Yoshii, & Sonehara, 2016) .

In addition, developing new products or optimizing existing products is also an application of Big Data technology in tourism industry. An optimization model was fitted for cruise route with the demand features as predictors, including navigation time, port of call, and so on to minimize operation cost meanwhile satisfy the demand of consumers (Ou, 2015). The analysis of mobile track data provided by the mobile phone operator could provide information of tourist flow behavior to inspire tourism product development (Lu & Zhong, 2016) .

Despite the applications “behind the stage”, there are also applications that could be reached by the customers to improve their travel experience. A study was conducted on college students to explore the impact of Location Based Services (LBS) on the level of satisfaction of travel experience (Tu, Li, Yu, Zhou, & Yang, 2015) . The result of the survey supported such impact. In another study published in 2007, 421 tourists were experimented to explore a city in Germany with mobile tour recommender systems. These tourists were found

discovering four times more sights and stayed there twice as long as the control group did (Modsching, Kramer, Hagen, & Gretzel, 2007) .

PART THREE

Introduction

In this section, a case will be introduced first on applications of Big Data Technology in tourism industry. Then the result of systematic literature review will be given out. Last but not least, managerial implications and future study will be discussed.

Case Study

Retrieving data from community websites, including photo sharing ones and view/opinion ones, researchers conducted a study to investigate how much new knowledge could be transferred from those data and then published it on 2014 IEEE International Conference on Big Data, trying to prove that such data and analysis do provide new knowledge (Chareyron, Da-Rugna, & Raimbault, 2014) . Table 5 summarizes the data resource.

Table 5

Photography Sharing Website Facts

	Panoramio	Flickr	Instagram
Photos	> 60M	> 170M	> 180M
Users	> 2.8M	> 1.7M	> 13M
Image metadata	GPS data； Camera timestamp； Tags, Title, Exif data		Location； Timestamp； Title, Tags
User information	Group Comments	Group; Comments Location	Comments

Note. Retrieved from “Big data: A new challenge for tourism”, by Chareyron, G., Da-Rugna, J., & Raimbault, T., 2014, Paper presented at the *2014 IEEE International Conference on Big Data*, 5-7.

By new knowledge, they mean something different from classical output such as the golden bridge is the most popular site for San Francisco. Through examining typical and atypical behaviors to form communities, such analysis would help business runners better comprehend and then cater to customers’ need, which could not be provided by traditional mainstream analysis.

Researchers first built up a virtual image of photo density in Paris, as shown in Figure 3. Although simple and direct, the visualization of raw data unveiled that photos taken and posted on the websites checked by tourists were concentrated to certain sites instead of evenly distributed. With red representing higher photo density and green for lower density, this map could tell the places where tourists inhabited frequently, which could be totally distinct from normal geographic space. Basically, the redder the site is on map, the more photographed it is by tourists, further indicating that the more visited it is by the same group of tourists.

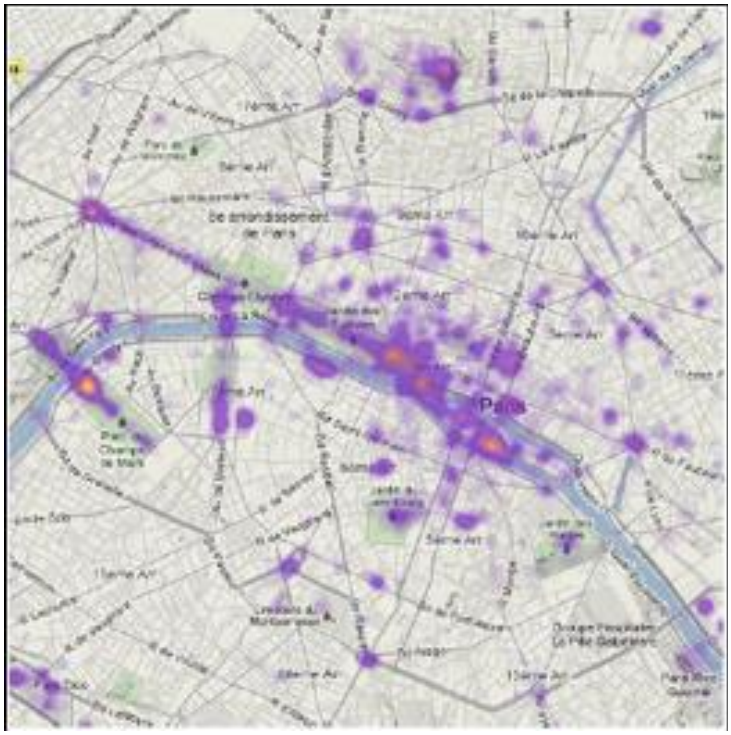

Figure 3. Photo Density of Paris. Reprinted from "Big data: A new challenge for tourism", by Chareyron, G., Da-Rugna, J., & Raimbault, T., 2014, Paper presented at the *2014 IEEE International Conference on Big Data*, 5-7. doi:10.1109/BigData.2014.7004475

Some photos were also "geotagged" and ordered chronologically, making it possible to reconstruct the travelling path of each individual. After combining each individual route, the main routes could be discovered in a certain area, as is shown in Figure 4. Before this technology, the shortest path model was used to estimate and forecast tourists travel plan. While it was not always the case for tourists to take the shortest path, this model might not be accurate enough. If further divide the visitors into various groups and detect their path using similar method, operators could investigate the travelling behavior of a certain community, including which path they would prefer, which kind of photo they would take, and what kind of activities they would perform. Using such information, new tourism products could be designed and promoted to its target market.

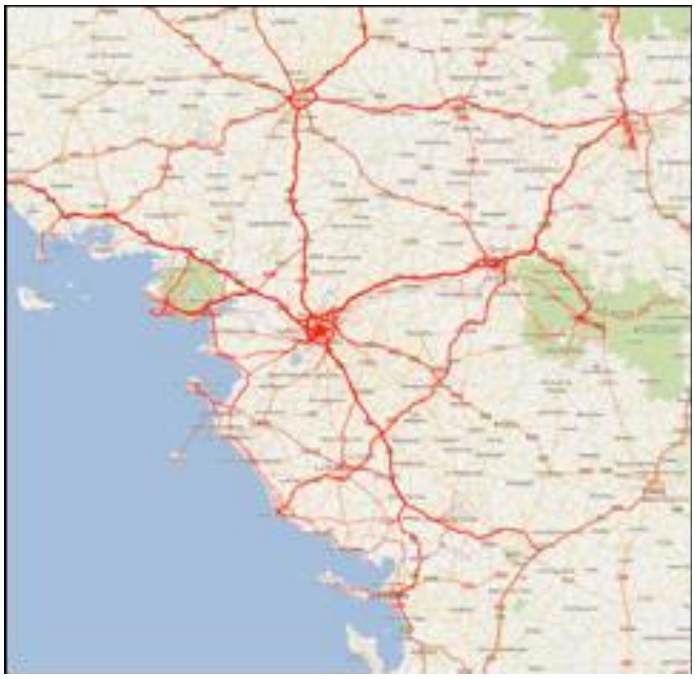

Figure 4. Extract Paths Overlay. Reprinted from “Big data: A new challenge for tourism”, by Chareyron, G., Da-Rugna, J., & Raimbault, T., 2014, Paper presented at the *2014 IEEE International Conference on Big Data*, 5-7.

Further more, these data could also be analyzed for characterization and detection of communities to examine the impact of community on the travel behavior of a community membership. For instance, we know Chinese prefer Chinese food even in a foreign country, but do Italian tend to choose Italian food in foreign countries as Chinese do? Questions like this could be answered through data mining to find whether there is a correlation between the characteristics of users, such as their age, sex, origin or even the model of their cameras, and the reviews or travel paths they chose. Then the communities could be recognized from similar behavior.

Last but not least, researchers also used data to assess the quality of reviews. As is known, the entities on review and opinion websites are often in contest. Potential customers may choose a restaurant, a hotel, or even a barber’s shop from candidates on the website based on

the scores and reviews given by other users. If so, there might be motivations to manipulate such scores, reviews and consequently rankings to draw the attention of potential customers. Comparing data of TripAdvisor and Hotel.com, researchers found that those rating and reviews were accurate in general, unless both websites were manipulated in the same way at the same time. They held the belief that quality assessment mechanics should be utilized to isolate fake opinion, unrepresentative photo, miss-geocoded data, and recurrent posting, etc.

Result

In the literature review, the definition of Big Data and its brief history were examined, through which it could be concluded that Big Data has been booming with a not so long history and still remains controversial in some area and deserves following studies. The intersection of Big Data and tourism industry is even younger – no articles were found in IEEE Xplorer Digital Library published before 2006. From this one could reasonably infer that Big Data did not draw the attention of tourism research until 2006, and that only after 2015 did this niche explode and become the “have-to-mention” topic in all forums within the tourism industry. The afore-introduced case has shown clearly how a tourism operator could collect data, use various technologies to dig out new knowledge from reachable data and then make business decisions.

Hot topics among the articles retrieved from IEEE Xplorer Digital Library included using Big Data to make predictions, support real-time decision-making process and develop and optimize tourism products. Last but not least, Big Data technology were also proposed to function as a piece of travelling tool to assist customers improve their travel experience, such as providing tour information based on the current location of a customer.

Conclusion

As conclusion, managerial implications drawn from the literature review will be presented and the limitations and the possible future study field will be discussed.

Managerial Implications

Just as SAS put it, it is not the amount of data that is important; it is what the organizations do with the data that matters. Above all the discussion of pros and cons of implementing Big Data practice, organizations should take the first step to admit the robustness of Big Data. Every day, the industry itself is generating a high volume of data, inclusive of customers' personal information and their purchasing behavior. Additionally, large amount of online UGC are generated by customers, from which their preference and personal views towards certain products, consuming behaviors, lifestyle and living patterns could be investigated. For decades marketers have been using marketing research tool such as focus group, deep interview, survey and on site observation to study consuming behavior, yet still may not be able to reach the whole target market because of the confine of time, budget and human resource. It is also a challenge to improve the validity of the result market research, eliminating the impact of researchers' interventions. With the implementation of Big Data, the process to reach the target market is simplified and the raw data collected are less influenced by the study procedure itself. It would be a pity that tourism operator step back and leave the data unattended, which will definitely bring forth opportunities to direct competitors. Implementing Big Data management is a long-term investment. It may require large expenditures at the very beginning, but the direct profit may not be apparent. This prevents some traditional business runners from investing on Big Data. But as a member of

the management team, one should evaluate an investment in the long run so that the company could function strategically in the fierce competition.

Then, how to take advantage of the abundant data is another consideration. In this paper, several Big Data Analytics are introduced and organizations may choose one or a combination of those analytics to do data mining based on the raw data available to them and their needs of information to make business decision. For example, a dynamic pricing service provider for airlines may track the time a customer log on the website, retrieve the weather information of the chosen travelling date from a third party, check the seasonal and holiday information of the date, trace back to the previous transactions of the same customer or similar characterized customers and then use predictive analytics to calculate an acceptable price for the customer in real-time, in the meantime maximize the potential profit from the current transaction. In contrast, a bank may use an audio analytics and interactive system to help customers make a payment or activate their new cards to save labor cost. Here, the properness outweighs the advancement of the techniques. Some analytics are perfect for some business but not others for a limited access to certain data. Some analytics might be robust but also costly, unaffordable to micro companies or start-ups. Some analytics may embrace a cost larger than projected revenue it could bring, and hence it makes no sense to continue with it for the present. In other word, when considering whether or not implementing Big Data and which one to use, the cost of acquiring and analyzing data and taking measures based on the analysis should be compared with the projected revenue it may bring to the table, and of course, both the direct and indirect benefits should be taken into considerations.

The last thing that deserves managers' focus is the ethics. Unless the wrongs the company has done will never be uncovered, the managers should always be prepared to handle such crisis. Amazon.com has been a benchmark for utilizing data mining to enlarge revenues, yet it also receives criticism from multiple sources for price differentiation, or even discrimination where they charged higher rates for loyal customers to make more money and lower rates to attract potential customers. This is reasonable for a company to optimize revenue, but may not be acceptable for loyal customers who are expecting some discount for being loyal to the website, not any punishment. Other ethic issue could be, should a review website delete negative remarks for an entity in exchange for profit? Should a property pay for popular bloggers and reviewers for their praise as a soft advertisement to bring in additional business? Once again, such actions have pros and cons. Managers should think over the result of disposure to the public before determination.

Limitations and Future Study

Up to now, only the strength and advantage of Big Data have been touched in this paper, which may make it too good to be true. Nevertheless, this worry is redundant as the term Big Data does have some drawbacks challenging scholars and data scientists.

First of all, data could lie. In his article, Schmidt (2010) suggested that data could lie to researchers, leading to false conclusions due to the way they were typically acquired, analyzed and interpreted. Tracing back to the notorious event "Dewey defeats Truman", as is shown in Figure 5, the sampling method should be found guilty for the misleading prediction. After conducting a poll using "quota sampling", Gallup predicted that Dewey would beat Truman with vote rates of 50% to 44%. Yet the actual result reversed: 50% were voted for

Truman and 44% for Dewey. In reality, quota sampling is defective in acquiring representative samples because no matter how carefully the sampling procedure is designed, human interventions in selecting samples are always biased. Except for the sampling error, another artifact of data analytics could be measurement error, which exists in all data due to the absence of perfect measures. For future study, research could be conducted to examine how Big Data could mislead the business process and how to avoid it to increase the reliability of the result of Big Data analysis and minimize the related risk.

Figure 5. Dewey Defeats Truman

Second, the lack of efficient statistical tool to process high volume of data instantaneously gives birth to the need of new statistical tool. Statistical methods are the foundation of predictive analytics. Traditional statistical methods are grounded from statistical significance, where a small sample is retrieved from the whole population to compare the significance of a particular relationship and then generalized to the original population. On the contrary, samples of Big Data Analytics are usually massive and able to represent the entire population. Under such circumstance, statistical significance may not be

applicable (Gandomi & Haider, 2015). New statistical tool customized for Big Data, or even customized for tourism industry is wanted so that future study should be conducted.

Lastly, the fakeness issue and the quality of data should also be taken into considerations. Again, the conclusion drawn from Big Data cannot be counted on if the data set is full of fakeness and flaw, otherwise the business decision might lead to a disaster. Yet the real situation is, fakeness in everywhere, especially online. Even if the online review site itself does not provide the service of "pay for high rankings", a lot of tour operators have started to "pay" their customers in exchange for positive reviews. It's not rare in some restaurants where a customer is encouraged to post positive reviews online by a 5% discount or a free drink. Under this circumstance, are the reviews still reliable? Or to what extend are these review reliable? Besides, unsatisfied customers are more tentative to complain their service failure on a review website than satisfied customers talking about their delightful experience. This may give rise to bias. So future study is needed on how to clean data deal with such fakeness and bias of collectable data.

References

Albusaidi, H. S., Udupi, P. K., & Dattana, V. (2016). Integrated data analytic tourism dashboard (IDATD). Paper presented at the *2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO),* 497-500. doi:10.1109/ICRITO.2016.7785006

Barik, R. K. , Das, P. K., & Lenka, R. K. (2016). Development and implementation of SOA based SDI model for tourism information infrastructure management web services. Paper presented at the *2016 6th International Conference - Cloud System and Big Data Engineering (Confluence),* 748-753. doi:10.1109/CONFLUENCE.2016.7723633

Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2015). An overview of big data opportunities, applications and tools. Paper presented at the *2015 Intelligent Systems and Computer Vision (ISCV),* 1-6. doi:10.1109/ISACV.2015.7105553

Bin Rusli, M. E., Yusof, A. M., & Ibrahim, N. (2006). Reservation through image visualization. Paper presented at the *International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06),* 38-42. doi:10.1109/CGIV.2006.77

Boodnah, K. D., Armoogum, V., Jaunky, K. D. ,& Armoogum, S. (2016). Towards smart tourism: An individual appreciation of porlwi-by-light festival. Paper presented at the *2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech),* 323-328. doi:10.1109/EmergiTech.2016.7737360

Buyya, R., Calheiros, R. N., & Dastjerdi, A. V., (2016). *Big data: Principles and paradigms.* Saint Louis, Missouri: Elsevier Science. Retrieved from https://ebookcentral.proquest.com/lib/unlv/detail.action?docID=4544734

Cacho, A., Estaregue, D., Figueredo, M., Lucas, J., Aurélio, M., Farias, H. . . . Alves, C. (2015). A smart destination initiative: The case of a 2014 FIFA world cup host city. Paper presented at the *2015 IEEE First International Smart Cities Conference (ISC2),* 1-6. doi:10.1109/ISC2.2015.7366223

Chareyron, G., Branchet, B.,& Jacquot, S. (2015). A new area tourist ranking method. Paper presented at the *2015 IEEE International Conference on Big Data (Big Data),* 2930-2932. doi:10.1109/BigData.2015.7364126

Chareyron, G., Da-Rugna, J., & Raimbault, T. (2014). Big data: A new challenge for tourism. Paper presented at the *2014 IEEE International Conference on Big Data*, 5-7. doi:10.1109/BigData.2014.7004475

Chen, X., & Zhou, L. (2015). Design and implementation of an intelligent system for tourist routes recommendation based on hadoop. Paper presented at the *2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS),* 774-778. doi:10.1109/ICSESS.2015.7339171

Cui, F., Lin, D., & Huang Y. (2015). The impact of perceived security on consumer E-loyalty: A study of online tourism purchasing. Paper presented at the *2015 IEEE First International Conference on Big Data Computing Service and Applications,* 314-322. doi:10.1109/BigDataService.2015.41

Duan, Y., Wang, Y., Wei, J., Kattepur, A., & Du, W. (2013). Constructing E-tourism platform based on service value broker: A knowledge management perspective. Paper presented at the *2013 IEEE International Conference on Big Data,* 17-24. doi:10.1109/BigData.2013.6691778

Fang, Y., Zhong, J., Liu, Y., & Guo, M. (2016). Semantic description and link construction of smart tourism linked data based on big data. Paper presented at the *2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA),* 32-36. doi:10.1109/ICCCBDA.2016.7529530

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

Guo, W., Liang, L., & Chen, C. (2010). Design for exhibition system based on RFID. Paper presented at the *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM),* 1-4. doi:10.1109/WICOM.2010.5601085

Guo, X. (2016). Application of meteorological big data. Paper presented at the *2016 16th International Symposium on Communications and Information Technologies (ISCIT),* 273-279. doi:10.1109/ISCIT.2016.7751635

Höpken, W., & Fuchs, M. (2016). Introduction: Special issue on business intelligence and big data in the travel and tourism domain. *Information Technology & Tourism, 16*(1), 1-4. doi:10.1007/s40558-016-0054-3

Hsu, L. Y. (2016). Inhering reliability in "signature" architecture: Developing remote control with effective order and efficient rules. Paper presented at the *2016 International Symposium on Computer, Consumer and Control (IS3C),* 235-238. doi:10.1109/IS3C.2016.70

Ichifuji, Y., Matsuo, Y., Koide, N., Akashi, N., Terai, Y., & Kobayashi, T. (2016). A study for understanding of tourist person trip pattern based on log data of wi-fi access points. Paper

presented at the *2016 IEEE International Conference on Big Data (Big Data),* 2167-2174. doi:10.1109/BigData.2016.7840846

Irudeen, R., & Samaraweera, S. (2013). Big data solution for sri lankan development: A case study from travel and tourism. Paper presented at the *2013 International Conference on Advances in ICT for Emerging Regions (ICTer),* 207-216. doi:10.1109/ICTer.2013.6761180

Jiang, M., Liu, H., & Niu, L. (2015). The evaluation studies of regional transportation accessibility based on intelligent transportation system: Take the example in yunnan province of china. Paper presented at the *2015 International Conference on Intelligent Transportation, Big Data and Smart City,* 862-865. doi:10.1109/ICITBS.2015.218

Khorasani, E. S., & Meybodi, S. M. N. (2010). A framework to improve E-tourism experience by using intelligent portal. Paper presented at the *2010 Seventh International Conference on Information Technology: New Generations,* 1226-1226. doi:10.1109/ITNG.2010.155

Kit, D., Kong, Y, & Fu, Y. (2016). Efficient image geotagging using large databases. *IEEE Transactions on Big Data, 2*(4), 325-338. doi:10.1109/TBDATA.2016.2600564

Koide, N., Ichifuji, Y., Yoshii, H., & Sonehara, N. (2016). Estimation of national tourism statistics based on wi-fi association log data. Paper presented at the *2016 IEEE International Conference on Big Data (Big Data),* 2175-2179. doi:10.1109/BigData.2016.7840847

Kuldeep, Garg, P. K., & Garg, R. D. (2016). Geospatial techniques for flood inundation mapping. Paper presented at the *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS),* 4387-4390. doi:10.1109/IGARSS.2016.7730143

Li, M., Bao, Z., Song, L., & Duh, H. (2016). Social-aware visualized exploration of tourist behaviours. Paper presented at the *2016 International Conference on Big Data and Smart Computing (BigComp),* 289-292. doi:10.1109/BIGCOMP.2016.7425930

Lo, C. C., Cheng, D. Y., Chen, C. H., & Yan, J. S. (2008). Design and implementation of situation-aware medical tourism service search system. Paper presented at the *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing,* 1-5. doi:10.1109/WiCom.2008.2523

Lu, D. D., & Zhong,Y. D. (2016). A tourist flows analysis system based on phone big data. Paper presented at the *2016 IEEE International Conference on Big Data Analysis (ICBDA),* 1-5. doi:10.1109/ICBDA.2016.7509822

Modsching, M., Kramer, R., Hagen, K. T., & Gretzel, U. (2007). Effectiveness of mobile recommender systems for tourist destinations: A user evaluation. Paper presented at the *2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications,* 663-668. doi:10.1109/IDAACS.2007.4488505

Mossucca, L., Goga, K., Spoto, G., Bolognesi, T., & Caragnano, G. (2016). BUSSOLA: A cloud collaborative platform of oriented services to passengers. Paper presented at the *2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS),* 420-425. doi:10.1109/CISIS.2016.91

Nakagawa, S., Akutsu, H., Tsuchiya, Y., Matsuhira, N., & Narita, M. (2016). Demonstration experiments of a robot service of stamp-rally and questionnaires for tourism destination marketing. Paper presented at the *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI),* 914-919. doi:10.1109/IIAI-AAI.2016.214

Ou, Y. (2015). Cruise tourism product developing model optimization based on demand forecasting. Paper presented at the *2015 International Conference on Intelligent Transportation, Big Data and Smart City,* 240-243. doi:10.1109/ICITBS.2015.66

Petroni, A., Biagi, M., Colonnese, S., Cusani, R., & Scarano, G. (2015). Vessels traffic estimation through image processing applied to acquisitions by hydrophones. Paper presented at the *OCEANS 2015 - Genova,* 1-4. doi:10.1109/OCEANS-Genova.2015.7271693

Press, G. (2013). A very short history of big data. *Forbes.* Retrieved from http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#4c92647155da

Sato, A. H. (2015). Microdata analysis of the accommodation survey in japanese tourism statistics. Paper presented at the *2015 IEEE International Conference on Big Data (Big Data),* 2700-2708. doi:10.1109/BigData.2015.7364070

Schmidt, F. (2010). Detecting and correcting the lies that data tell. *Perspectives on Psychological Science, 5*(3), 233-242. doi:10.1177/1745691610369339

Shafiee, S., & Ghatari, A. R. (2016). Big data in tourism industry. Paper presented at the *2016 10th International Conference on E-Commerce in Developing Countries: With Focus on E-Tourism (ECDC),* 1-7. doi:10.1109/ECDC.2016.7492979

Shia, B. C., Li, S. C., Chen, C. J., & Hsu, S. H. (2009). A market survey and prediction of lottery industry of taiwan. Paper presented at the *2009 International Conference on New Trends in Information and Service Science,* 853-856. doi:10.1109/NISS.2009.235

Sun, Y., Song, H., Jara, A. J., &Bie, R. (2016). Internet of things and big data analytics for smart and connected communities. *IEEE Access, 4,* 766-773. doi:10.1109/ACCESS.2016.2529723

Tsuda, H., Ando, M., & Ichifuji Y. (2016). Hotel plan popularity factor analysis of hotels in the keihanshin region. Paper presented at the *2016 IEEE International Conference on Big Data (Big Data)*, 2217-2224. doi:10.1109/BigData.2016.7840852

Tu, T., Li, Y., Yu, Y., Zhou, Z., & Yang, X. (2015). Research on LBS impact on college students' travel experience. Paper presented at the *2015 IEEE First International Conference on Big Data Computing Service and Applications*, 406-411. doi:10.1109/BigDataService.2015.18

Wang, J., Wang, C., & Wang, X. (2010). An experimental research on fusion algorithms of ETM+ image. Paper presented at the *2010 18th International Conference on Geoinformatics*, 1-6. doi:10.1109/GEOINFORMATICS.2010.5567950

Vatanasakdakul, S., & Aoun, C. (2009). Social structures, isomorphic pressures, and B2B utilisation in the thai tourism industry. Paper presented at the *2009 42nd Hawaii International Conference on System Sciences*, 1-10. doi:10.1109/HICSS.2009.397

Xiong, C., Niu, L., Liang, X., & Jiang, M. (2015). The trend analysis of the second industrial structure changes in yunnan province of china based on the big data technology and dynamic shift-share method analysis. Paper presented at the *2015 International Conference on Intelligent Transportation, Big Data and Smart City*, 939-942. doi:10.1109/ICITBS.2015.238

Xiong, C., & Wu, Y. (2015). The geographic environment analysis of regional economic development of yunnan province of china based on the big data technology. Paper presented at the *2015 International Conference on Intelligent Transportation, Big Data and Smart City*, 869-872. doi:10.1109/ICITBS.2015.220

Yang, R. (2016). Design of an information system for smart scenic spots. Paper presented at the *2016 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, 328-331. doi:10.1109/ICSGEA.2016.57

Yuan, Y., Du, J., Fan, D., & Lee, J. (2016). Design and implementation of tourism activity recognition and discovery system. Paper presented at the *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, 781-786. doi:10.1109/WCICA.2016.7578321

Zhang, Y., Shu, S., Ji, Z., & Wang, Y. (2015). A study of the commercial application of big data of the international hotel group in china: Based on the case study of marriott international. Paper presented at the *2015 IEEE First International Conference on Big Data Computing Service and Applications*, 412-417. doi:10.1109/BigDataService.2015.59

Zou, L., Liu, Q., Zhang, C., & Yang, H. (2016). An approach to applying creative computing in tourism by constructing a big data based knowledge system framework. Paper presented at the *2016 22nd International Conference on Automation and Computing (ICAC)*, 244-249. doi:10.1109/IConAC.2016.7604926

Appendix A

Full Article List for Systematic Literature Review

Document Title	Authors
A Market Survey and Prediction of Lottery Industry of Taiwan	Sato, 2015.
A new area tourist ranking method	Chareyron, et al, 2015.
A smart destination initiative: The case of a 2014 FIFA world cup host city	Cacho, et al, 2015.
A study for understanding of tourist person trip pattern based on log data of Wi-Fi access points	Ichifuji, et al, 2016.
A Study of the Commercial Application of Big Data of the International Hotel Group in China: Based on the Case Study of Marriott International	Zhang, Shu, Ji, & Wang, 2015.
An overview of big data opportunities, applications and tools	Benjelloun, Lahcen, & Belfkih, 2015.
Application of meteorological big data	Guo, 2016.
Big data in tourism industry	Shafiee & Ghatari, 2016.
Big data solution for Sri Lankan development: A case study from travel and tourism	Irudeen, & Samaraweera, 2013.
Big data: A new challenge for tourism	Chareyron, et al, 2014.
Cruise Tourism Product Developing Model Optimization Based on Demand Forecasting	Ou, 2015.
Effectiveness of Mobile Recommender Systems for Tourist Destinations: A User Evaluation	Modsching, et al, 2007.
Estimation of national tourism statistics based on Wi-Fi association log data	Koide, et al, 2016.
Hotel plan popularity factor analysis of hotels in the Keihanshin region	Tsuda, et al, 2016.
Microdata analysis of the accommodation survey in Japanese tourism statistics	Sato, 2015.
Research on LBS Impact on College Students' Travel Experience	Tu, et al, 2015.
Social Structures, Isomorphic Pressures, and B2B Utilisation in the Thai Tourism Industry	Vatanasakdakul & Aoun, 2009.
The Geographic Environment Analysis of Regional Economic Development of Yunnan Province of China Based on the Big Data Technology	Xiong & Wu, 2015.
The Impact of Perceived Security on Consumer E-Loyalty: A Study of Online Tourism Purchasing	Cui, Lin, & Huang, 2015.
The Trend Analysis of the Second Industrial Structure Changes in Yunnan Province of China Based on the Big Data Technology and Dynamic Shift-Share	Xiong, Niu, Liang, & Jiang, 2015.

Document Title	Authors
Method Analysis	
Towards Smart Tourism: An individual appreciation of Porlwi-By-Light festival	Boodnah, Armoogum, Jaunky, & Armoogum, 2016.
A Framework to Improve E-Tourism Experience by Using Intelligent Portal	Khorasani, & Meybodi, 2010.
A tourist flows analysis system based on phone big data	Lu, & Zhong, 2016.
An approach to applying creative computing in tourism by constructing a Big Data based Knowledge System Framework	Zou, Liu, Zhang & Yang, 2016.
An experimental research on fusion algorithms of ETM+ image	Wang, Wang, & Wang, 2010.
BUSSOLA: A Cloud Collaborative Platform of Oriented Services to Passengers	Mossucca, Goga, Spoto, Bolognesi, & Caragnano, 2016.
Constructing E-Tourism platform based on service value broker: A knowledge management perspective	Duan, et al, 2013.
Demonstration Experiments of a Robot Service of Stamp-Rally and Questionnaires for Tourism Destination Marketing	Nakagawa., Akutsu, Tsuchiya, Matsuhira, & Narita, 2016.
Design and implementation of an intelligent system for tourist routes recommendation based on Hadoop	Chen & Zhou, 2015.
Design and Implementation of Situation-Aware Medical Tourism Service Search System	Lo, Cheng, Chen, & Yan, 2008.
Design and implementation of tourism activity recognition and discovery system	Yuan, Du, Fan, & Lee, 2016.
Design for Exhibition System Based on RFID	Guo, Liang, & Chen, 2010.
Design of an Information System for Smart Scenic Spots	Yang, 2016.
Development and implementation of SOA based SDI model for tourism information infrastructure management web services	Barik, Das, & Lenka, 2016.
Efficient Image Geotagging Using Large Databases	Kit, Kong, & Fu, 2016.
Geospatial techniques for flood inundation mapping	Kuldeep, Garg, & Garg, 2016.
Inhering Reliability in "Signature" Architecture: Developing Remote Control with Effective Order and Efficient Rules	Hsu, 2016.
Integrated data analytic tourism dashboard (IDATD)	Albusaidi, Udupi, & Dattana, 2016.
Internet of Things and Big Data Analytics for Smart and Connected Communities	Sun, Song, Jara, &Bie, 2016.
Reservation Through Image Visualization	Bin Rusli, Yusof, & Ibrahim, 2006.

Document Title	Authors
Semantic description and link construction of smart tourism linked data based on big data	Fang, Zhong, Liu, & Guo, 2016.
Social-aware visualized exploration of tourist behaviours	Li, Bao, Song, & Duh, 2016.
The Evaluation Studies of Regional Transportation Accessibility Based on Intelligent Transportation System: Take the Example in Yunnan Province of China	Jiang, Liu, & Niu, 2015.
Vessels traffic estimation through image processing applied to acquisitions by hydrophones	Petroni, Biagi, Colonnese, Cusani, & Scarano, 2015.