A bibliometric analysis on Artificial intelligence in Tourism. 
State of the art and future research avenues

MARTINA NANNELLI, FRANCESCO CAPONE, LUCIANA LAZZERETTI

Working Paper N. 3/2022

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

The findings, interpretations, and conclusions expressed in the working paper series are those of the authors alone. They do not represent the view of Dipartimento di Scienze per l'Economia e l'Impresa.
A bibliometric analysis on Artificial intelligence in Tourism.  
State of the art and future research avenues  

Nannelli M.¹, Capone F.², Lazzeretti L.³

Abstract

The tourism industry has undergone a deep transformation driven by the information and communication technologies (ICTs) development that has taken hold in recent years thanks to innovations in Artificial Intelligence (AI) and its tools.

Although a growing literature is being produced on this topic, scientific research on AI and tourism is still fuzzy and fragmented. The aim of this work is to explore the current state of art and possible future developments of Artificial Intelligence and its tools in the field of tourism.

Different methodologies were combined to achieve this objective. The study develops first a bibliometric analysis using the ISI web database and applies then social network analysis (SNA) to identify the main authors and to explore the intellectual structure of this field through keyword co-occurrence, then a qualitative literature review has been developed to investigate the main research themes, applications and developments.

The findings confirm the industry’s direction towards digitization and robotization of services and identify some of the main research strands: the use of Big Data (BD) for demand forecasting and customer satisfaction measurement; Augmented Reality (AR) and Virtual Reality (VR) experience for value co-creation processes; the Covid-19 pandemic, healthcare and social distances issues and service robots; and finally, the smart tourism trends.

Keyword: artificial intelligence; tourism industry; bibliometric analysis; social network analysis; literature review.

JEL Codes: L83, Z30, L86

¹ Corresponding Author. University of Florence, Department of Economics and Management, Via delle Pandette 9, 50127 Florence, Italy. E-mail: martina.nannelli@unifi.it
² University of Florence, Department of Economics and Management, Via delle Pandette 9, 50127 Florence, Italy. E-mail: francesco.capone@unifi.it
³ University of Florence, Department of Economics and Management, Via delle Pandette 9, 50127 Florence E-mail: luciana.lazzeretti@unifi.it
1. Introduction

Over the past two decades, the tourism industry has undergone a deep transformation driven by ICTs development that had a disruptive impact on structures, processes and practices (Buhalis and Law, 2008; Sigala et al., 2015; Sigala, 2018). As Buhalis (2019) underlines, the development of AI’s tools, such as Machine Learning, Deep Learning and Data Mining, is among the main reasons for the current expansion and enhancement of the industry, especially in the field of robotic, virtual and AR and conversational software. While access to BD sources in the tourism industry is expanding the network of information for increasingly accurate demand forecasting and consumer profiling.

Nowadays, due to the Covid-19 pandemic, the tourism industry has suffered a setback which is making necessary to rethink the industry to promote a post-pandemic restart in which AI technologies can be the engine for a new renaissance of the whole industry (Hall et al., 2020).

Studies on AI and its tools are increasing in many fields and disciplines. Lazzeretti (2020) widely discusses digital transformation in the algorithmic society through a cultural and creative approach. Nevertheless the tourism industry is making extensive use of AI tools, studies on the impact of AI and its tools in this field are still very few (Buhalis et al., 2019; Mariani, 2020; Lv et al., 2021; Mariani and Baggio, 2021).

The aim of this work is to explore the current state of the art and possible future developments on the topic of AI and its tools in field of tourism.

We combine different approaches to achieve our aim. The study follows a bibliometric approach, already successfully applied in social science disciplines, to collect data and analyse the spread and distribution of scientific publications (Lazzeretti et al., 2014; Sedita et al., 2020; ), with a review of the literature (Lazzeretti et al., 2022). We collect papers during the period 1986-2021 using WOS by combining keywords that identify AI (i.e. artificial intelligence, big data, machine learning, data mining, etc.) with tourism topics (i.e., tourism, destination, hospitality, visitors, etc) chosen through to the information from the literature and contribution from experts. The keywords are searched in business and economics (e.g., business, economics, geography, management, social issue and urban studies) and tourism WOS categories (e.g. Hospitality, leisure, sport tourism).

We then apply SNA to investigate co-authorship networks and keywords co-occurrence in the articles to map the intellectual structure of AI in tourism.

Finally, we select the most important articles based on the higher average citation per year, to conduct a qualitative literature review to go deep into understanding the main themes emerging from the topic and their trends. This allows us to elaborate and discuss new concepts.
The study shows that the scientific publications on AI for the tourism disciplines have drastically increased in the last few years meaning that AI is becoming more central in the field. Moreover, the study identifies the main emerging topics, namely: the use of BD for demand forecasting and customer satisfaction measurement; the AR and VR experience for value co-creation process; the Covid-19 pandemic and the connection with service robots; finally, the smart tourism trends.

The article is structured as follows. Section 2 provides an overview of the researches on AI and its tools in the tourism industry. Section 3 explains the methodology. Section 4 presents the results of the bibliometric analysis. Section 5 discusses the relevant topics emerging from the qualitative literature review. Finally, conclusions draw suggestions for future research that emerged from the analysis of the literature.

2. The previous research on AI in the tourism literature

Technological development has driven the tourism industry to undergo a profound revolution since the early 1960s, while the rise of the Internet in the late 90s triggered the greatest change by producing a disruption of all sources of competitive advantage (Buhalis and Law, 2008). Many innovative transformations have taken place over time: form of global distribution systems (Buhalis and Licata, 2002); booking and reservation systems (Ghose et al., 2012); social media and mobile applications (Sigala, 2018). In particular, what has given the definite push towards technological drift in recent years is the developments of AI innovations: new search engines and web data mining (Xiang et al., 2015); new tools to enhance and alter the tourist's experience such as AR and VR devices (Guttentag, 2010); new tools for processing ever larger amounts of data such as Machine Learning, Deep Learning, Neural networks, and enabling increasingly accurate estimates and forecasts in real time (Buhalis and Sinatra, 2019; Law et al., 2019). Moreover, AI’s tools are motivating the automation of processes and services in the tourism industry being recently used as a way to interact directly with customers in frontline services through technological innovations related to robotics such as chatbots and human like robots (Belanche et al., 2021). This rapid evolution has made ICTs indispensable in the different stages of tourism supply and demand from marketing to distribution, promotion and coordination revolutionising both products and services and contributing to create a smart and connected tourism environment (Porter and Heppelmann, 2014; Gretzel et al., 2015).

The Covid-19 pandemic has given further impetus to the application of AI tools in the industry, halting tourist flows and forcing supplementary digitisation of the industry (Zeng et al., 2020). Several studies have been carried out on a wide range of topics linking the triad of AI, tourism industry and pandemic. A first common element emerging from several studies is the increasingly
pressing need to make accurate estimates and forecasts on the demand for tourism products and services, especially as a consequence of the Covid-19 pandemic. In this regard, an invaluable source of tourist data to be processed for demand forecasting employing advanced computational methodologies comes from social networks and sharing economy digital platforms, which are thus one of the most discussed topics in the connection between AI and tourism (Garcia-Palomares et al., 2015; Tussyadiah and Zach, 2017; Cheng et al., 2019). In addition, since the current pandemic has made it necessary to rethink the tourism industry through tools that allow to carry out essential services in compliance with the health and social distancing standards, related scientific studies also emerge. The application of robotics such as cleaning and sanitary robots, welcoming services robots, information totems, to tourism represent one of the most relevant innovations. In the same direction, AR and VR tools have also made it possible to enhance the tourist experience where the difficulties of travelling had made it impossible to experience places. For example, the possibility of guided tours directly from digital devices has emerged (Loureiro et al., 2020).

Some attempts of qualitative analysis have been conducted to draw up the state of the art of the industry and to trace the development of this extraordinary digital transformation (Buhalis and Law, 2008; Leung et al., 2013; Gretzel et al., 2015).

Recently, some bibliometrics has covered the binomial topic of tourism and new technologies. Benckendorff (2009) presents a bibliometric analysis of tourism literature identifying emerging themes and influences. One of the most interesting results concerns the emergence of the Internet as a core keyword in recent years, testifying that the tourism industry is moving towards a technological drift that sees its development driver precisely in the appearance and use of this technology. Gomezelj (2016) uses a bibliometric approach to explore innovation in tourism identifying nine co-citation networks, or clusters where innovation is playing a central role. Ali et al. (2019) conducted a bibliometric analysis covering 30 years of research. Most recent topics are customer experience and satisfaction and gaming and demand prediction and both are connected to the use of technologicaldigital tools linked to AI. Nusair (2020) employs a bibliometric analysis in hospitality and tourism journals, discovering the emergence of new trends such as BD, e-tourism, green experience and smart tourism and the appearance of digital platforms in the tourism field. Zhang et al. (2020) employ a bibliometric analysis shedding light on global tourism demand forecasting and techniques, while Bastidas-Manzano et al. (2021) using a bibliometric approach to map the development of smart tourism destinations find that among the main emerging topics there are ICT technology, AR, intelligent tourism ecosystem and the vital role covered by BD in the field.

The only contribution focused on AI and tourism is however the work of Lv et al. (2021). The authors have conducted a bibliometric analysis on AI and BD in hospitality and tourism by reviewing
270 articles published from 2007 to 2020. Findings reveal that, although BD and AI have gradually attracted attention in tourism, attempts are still fuzzy and only few studies have been conducted to update and summarize the evolution of new technologies in the industry. The authors agree that this line of research could serve as a guideline for scholars to understand the current state of research and develop new insights for the industry, both for private businesses to forecast and manage flows and for policymakers to develop a smart tourism environment. This interesting work focuses on tourism journals which were pre-selected upstream of the construction of the database. However, since the subject combines different fields of research both in social and hard sciences such as tourism - which is per se a multidisciplinary industry -, computer science, engineering, mathematics, it seems essential that a comprehensive analysis also considers journals not specialised in tourism. Furthermore, as shown by the findings of Mariani and Baggio (2021) who carried out a systematic literature review on BD and analytics, most scientific articles are published in academic journals whose main references area is not focused on tourism or hospitality.

These results, on the one hand, highlighted the growing trend of AI research and related topics in the field of tourism, the spread of new technological tools throughout the industry and their increasing necessity to overcome several limitations; on the other, they highlighted the need for more studies to support this research context which is still fuzzy and fragmented.

3. Data source and empirical methodology

The study aims to investigate the current state of the art and future research avenues of AI in the field of tourism research by apply, first, a bibliometric analysis using ISI web database and then SNA to identify the main authors and to explore the intellectual structure of the field through the keywords co-occurrence (Lazzeretti et al., 2014; Sedita et al., 2020). Finally, we develop a review of the literature to afford several wide and blurry topics emerging from the previous analysis. A combination of these approaches may contribute to better explore a wide topic that connects multidisciplinary fields of research considering a large number of studies and debating the topic deeply.

This study collects documents from the ISI - Web of Science (ISI WoS) database, one of the largest sources of quality scientific publications internationally (Sutherland and Jarrahi, 2018). We collect papers during the period 1986-2021 by combining keywords closely related to the concept of AI (i.e., big data, machine learning, data mining, etc.) with tourism topics (i.e., tourism, destination, hospitality, visitors, etc) chosen through a preliminary literature review. The following criteria have been applied to restrict the perimeter of the analysis consistent with the research objective:
- Language: to avoid language constraints, only English contributions have been selected being those that could be adequately analysed by the authors.

- Source: following previous literature, the selection has involved only peer-reviewed articles in scientific journals to ensure the quality of the sample and comparability with other studies.

- Research domain: To maintain a wide perspective on the issue, the selected contributions pertain to the general social sciences categories, for example Business, Economics, Hospitality Leisure Sport Tourism, Management, Social issue and Urban studies. These categories have been selected among the disciplines of ISI WoS database.

We limited the analysis to articles with at least 1 citation in order to exclude marginal researches and we collected 1,297 scientific papers. The papers were then manually screened, reading the title, abstract and if necessary the full paper, excluding articles that are inconsistent with the topics and aim. The final dataset counts 696 articles.

Finally, we combine the previously quantitative methods with a qualitative literature review. Qualitative literature reviews in a narrative style have been used in the field of tourism research (Seraphin et al. 2018) and, in our case, can contribute to tracing back the most discussed topics and help to draw an agenda for future research. To conduct the qualitative literature (Section 5), we select the most important articles based on the higher average citation per year (> 20).\(^4\) We refined our selection through an in-depth reading of abstracts and the entire article when needed. The literature review was carried out of the final sample of 52 articles plus 11 which were considered relevant to reach our aim, for a total of 63 scientific articles.

4. The role of artificial intelligence in Tourism

4.1. The evolution of publications on AI in tourism

Figure 1 offers an overview of the evolution of the topic over time. The number of articles on AI in the tourism literature has experienced growth over the years, as expected, reaching up to 166 articles per year in 2021. This result indicates how hot the topic is and the degree to which the scientific community is investigating it. This growth however is very recent in fact until 2010 very few articles are found. First publications start in 1995, no evidence of scientific articles on the subject is found before this date. Simultaneously with the growth in the number of articles, the number of

\(^4\) The literature often uses this method because the most frequently cited articles are those that have found the greatest consensus in the academy, and using average citation per year allows us not to underestimate more recent articles in favour of older ones (Lazzeretti et al, 2017).
citations also experiences a surge (FIG. 2). In 2020 citations reached 1.264 almost doubling in just two years.

**FIG. 1 THE EVOLUTION OF PUBLICATIONS OF AI IN TOURISM - 1995-2021**

![Graph showing the evolution of publications of AI in tourism from 1995 to 2021. Source: Authors’ elaboration based on ISI WoS database.]

**FIG. 2 EVOLUTION OF CITATION OF AI IN TOURISM - 1995-2021**

![Graph showing the evolution of citation of AI in tourism from 1995 to 2021. Source: Authors’ elaboration based on ISI WoS database.]

7
Another important result concerns the journals in which the contributions have been published (Table 1). Most of the articles are published in Tourism Management (105) which has more than twice the number of articles published in the second and third journals, namely International Journal of Contemporary Hospitality Management (47) and International Journal of Hospitality Management (44). These journals are related to management disciplines, emphasizing that these issues are primarily discussed in this field of research. In addition, many other journals among the most published journals are related to the disciplines of management and marketing, for example: Journal of Destination Marketing & Management (27), Tourism Management Perspectives (23), etc. It should be noted that there is a group of journals that, alongside the managerial area, deal with topics about Economics and Social Sciences. Among these for example: Current Issues in Tourism (31), Annals of Tourist Research (30), Tourism Review, Journal of Travel Research (19), etc. Others, however, place the research area of computer science and information systems alongside management-related topics in Journal of Hospitality and Tourism Technology (21), and Information Technology & Tourism (15). Another important element concerns the entry of journals dedicated to computer-engineering sciences. Among these, we have Expert Systems With Applications (18), Information Technology & Tourism (15), etc.

The search identifies multidisciplinary journals that focus mainly on tourism and new technologies, as was to be expected given the selection of keywords. However, not having chosen the journals a priori allowed us to show that the studies theme is analysis from different perspectives and disciplines.

**Table 1 Most publishing journals from 1986 to 2021**

<table>
<thead>
<tr>
<th>N.</th>
<th>Journal</th>
<th>N. of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TOURISM MANAGEMENT</td>
<td>105</td>
</tr>
<tr>
<td>2</td>
<td>INTERNATIONAL JOURNAL OF CONTEMPORARY HOSPITALITY MANAGEMENT</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>INTERNATIONAL JOURNAL OF HOSPITALITY MANAGEMENT</td>
<td>44</td>
</tr>
<tr>
<td>4</td>
<td>CURRENT ISSUES IN TOURISM</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>ANNALS OF TOURISM RESEARCH</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>JOURNAL OF DESTINATION MARKETING &amp; MANAGEMENT</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>TOURISM MANAGEMENT PERSPECTIVES</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>JOURNAL OF HOSPITALITY AND TOURISM TECHNOLOGY</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>JOURNAL OF TRAVEL RESEARCH</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>EXPERT SYSTEMS WITH APPLICATIONS</td>
<td>18</td>
</tr>
<tr>
<td>11</td>
<td>ISPRS INTERNATIONAL JOURNAL OF GEO- INFORMATION</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>JOURNAL OF TRAVEL &amp; TOURISM MARKETING</td>
<td>17</td>
</tr>
<tr>
<td>13</td>
<td>TOURISM ECONOMICS</td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td>JOURNAL OF HOSPITALITY MARKETING &amp; MANAGEMENT</td>
<td>16</td>
</tr>
<tr>
<td>15</td>
<td>INFORMATION TECHNOLOGY &amp; TOURISM</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>TOURISM REVIEW</td>
<td>11</td>
</tr>
<tr>
<td>17</td>
<td>JOURNAL OF HOSPITALITY AND TOURISM MANAGEMENT</td>
<td>11</td>
</tr>
</tbody>
</table>
4.2. The most publishing authors

There are 1,599 authors of the 696 articles for 2,138 total co-authorships. The average number of authors per article is 3 confirming the increasing trend of co-authorship in many disciplines. The network of authors is not very cohesive and hierarchical, in fact out of 1,599 authors, 1,337 publish only 1 paper. Only 99 authors publish at least 3 papers on AI and tourism and record 22% of all papers.

Table 2 shows the top 20 authors with at least 6 publications on the topic of analysis in the period. The table also shows the affiliations at the time of last publication. These 20 authors alone publish 10% of the total publications on tourism and AI. Delving deeper into the table we can identify the important authors such as Law, Gang, Dieck, Mariani, and Buhalis at the top and some globally spread schools with much presence in the UK, Australia, China, and the USA. Among the most important we can highlight Manchester Metropolitan University (Dieck, Jung, Jung), Deakin University (Huy, Rong), Bournemouth University (Buhalis), University of Reading (UK) (Mariani and Borghi), etc.

**TABLE 2: MOST PUBLISHING AUTHORS.**

<table>
<thead>
<tr>
<th>NR</th>
<th>Author</th>
<th>#</th>
<th>University*</th>
<th>Nationality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Law, Rob</td>
<td>38</td>
<td>University of Macau Development Foundation (UMDF)</td>
<td>China</td>
</tr>
<tr>
<td>2</td>
<td>Li, Gang</td>
<td>19</td>
<td>School of Hospitality and Tourism Management, Centre for Competitiveness of the Visitor Economy, Surrey</td>
<td>UK</td>
</tr>
<tr>
<td>3</td>
<td>Dieck, M. Claudia Tom</td>
<td>13</td>
<td>Creative AR &amp; VR Hub, Manchester Metropolitan University</td>
<td>UK</td>
</tr>
<tr>
<td>4</td>
<td>Mariani, Marcello M.</td>
<td>13</td>
<td>Henley Business School, University of Reading, Greenslands, Henley on Thames Oxfordshire and University of Bologna</td>
<td>UK</td>
</tr>
<tr>
<td>5</td>
<td>Ivanov, Stanislav</td>
<td>10</td>
<td>Varna University of Management, Varna</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>6</td>
<td>Buhalis, Dimitrios</td>
<td>9</td>
<td>International Centre for Tourism and Hospitality Research, Bournemouth University, Poole</td>
<td>UK</td>
</tr>
<tr>
<td>7</td>
<td>Huy Quan Vu</td>
<td>9</td>
<td>Department of Information Systems and Business Analytics, Deakin University</td>
<td>Australia</td>
</tr>
<tr>
<td>8</td>
<td>Moro, Sergio</td>
<td>8</td>
<td>Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Lisbon</td>
<td>Portugal</td>
</tr>
<tr>
<td>9</td>
<td>Song, Haiyan</td>
<td>8</td>
<td>Hospitality and Tourism Research Centre, School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Kowloon</td>
<td>China</td>
</tr>
<tr>
<td>10</td>
<td>Jung, Timothy Hyungsoo</td>
<td>7</td>
<td>Creative AR &amp; VR Hub, Manchester Metropolitan University</td>
<td>UK</td>
</tr>
<tr>
<td>11</td>
<td>Rita, Paulo</td>
<td>7</td>
<td>NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa</td>
<td>Portugal</td>
</tr>
<tr>
<td></td>
<td>Name</td>
<td>Affiliation</td>
<td>Country</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>--------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Borghi, Matteo</td>
<td>Department of Management, University of Reading, Reading</td>
<td>UK</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Flaviano, Carlos</td>
<td>Faculty of Economy and Business, University of Zaragoza,</td>
<td>Spain</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Jung, Timothy</td>
<td>Creative AR &amp; VR Hub, Manchester Metropolitan University</td>
<td>UK</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kirilenko, Andrei P.</td>
<td>Department of Tourism, Hospitality, and Event Management,</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>University of Florida</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Pan, Bing</td>
<td>Department of Recreation, Park, and Tourism Management,</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pennsylvania State University</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Park, Sangwon</td>
<td>Kyung Hee University</td>
<td>South Korea</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Rong, Jia</td>
<td>School of Information Technology, Deakin University</td>
<td>Australia</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Wang, Shouyang</td>
<td>CFS, MDIS, Academy of Mathematics and Systems Science, Chinese Academy of</td>
<td>China</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sciences, Beijing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Xiang, Zheng</td>
<td>The Howard Feiertag Department of Hospitality and Tourism Management,</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pamplin College of Business, Virginia Tech, Blacksburg, USA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: our elaboration. Affiliation at the latest articles.

Finally, it is also possible to analyse through SNA the co-authorship networks publishing in AI and tourism. Figure 3 presents the 189 top authors of the 696 publications. All the authors highlighted in Table 2 can be found in figure 3 identifying now collaborations and schools. From the graph it is possible to identify the main stable collaborations of relevant research groups. We can find in the centre of the graph authors like Law, Gang and Buhalis, Huy Quan Vu, etc., with a large sub-network of collaborations. To the right of the graph, we may identify a group of Portuguese authors including Sergio Moro, Paulo Rita, Ribeiro Ricardo, etc. At the bottom we find Dieck, Jung and Lee with several collaborations. In the middle Zheng together with his co-authors, and on the right Wang, etc.
4.3. Mapping the artificial intelligence landscape through the network of co-keywords

This section investigates and maps the intellectual structure of the field through the keyword co-occurrence analysis (Lazzeretti et al., 2017; Capone 2016; Beckendorff, 2009). With the help of SNA, it is possible to graphically represent the links among contributions and the most important keywords identified with the textual analysis. It was then possible to transform the matrix publication-keywords into a two-mode keywords-keywords graph. Fig. 4 represents the network of the most cited keywords providing information about which research trends are of most concern to researchers.

As might be expected, the most complex sub-network is composed of the Big Data keyword, which is the largest in size, central to the network and the one with the most links showing how BD studies are central to the field of tourism. BD are treated both from a theoretical point of view as a new field of research and empirically for the application of new quantitative methodologies. This duality is shown by the sub-networks that arise from the keyword BD. The first subnetwork concern connections between BD and the main BD analysis methodologies and consist of keyword deep
learning, big data analytics, sentiment analysis, and text mining. In addition to the main possibilities of BD in tourism for forecasting and management, eWom (electronic word of mouth), destination management, customer satisfaction, google trends, search query data, and tourism demand forecasting. The second sub-network is made up of the keyword BD linked to social media’s data sources which are digital platforms such as Tripadvisor, Airbnb, and Booking, and also linked to them is the term Sharing Economy which is the technological paradigm within which digital platforms were created. On the other, it is composed of the keywords that represent its potential for making predictions: satisfaction, customer satisfaction, visit intention, and tourist behaviour.

Another sub-network is represented by the keyword machine learning, a BD processing methodology which is linked to the keywords social media analytics, topic modelling, airbnb, behavioural intentions, geotagged photos, and content analysis. These words, as in the case of BD, partly refer to data processing methodologies, partly to data sources. Moreover, Machine Learning is linked to the keywords sentiment analysis and big data analytics, both representing BD methodologies.

A third sub-network is constituted by the keyword robot, which connects in a strong double thread (size of the connecting ties) with the word Artificial Intelligence, which is the focus of the research. The sub-network is then made up of other related words (automation, service automation, co-creation, human-robot interaction, service encounter, service robot, anthropomorphism, and finally hospitality, hotel hospitality and tourism) emphasizing and confirming the shift in the hospitality sector towards the robotization of services and the increasingly humanised characteristics of robots to be an integral part of the value co-creation process in the hospitality sector.

The fourth sub-network that emerges sees the words Virtual Reality and Augmented Reality as central nodes. Around them revolve the words destination image, destination marketing, market segmentation engagement, tourism marketing, immersion, presence, mental imagery, user experience, presence, technological embodiment, emotions, and interactivity. Also linked to this group of words is the Covid-19 pandemic. This network confirms the trend of tourism destinations towards immersive marketing that focuses on direct user experience. This group is important because it sets a new industry trend toward the virtual experience which is also a way to overcome the current pandemic showing how far the emergency has affected tourism.
5. Main research themes of AI in Tourism

This second part of the study applies a qualitative approach by analysing the literature to explore some of the emerging themes linking AI to the tourism context. In the following sub-paragraphs, we analysed the most cited articles (>20 average citation per year) for the literature review by reading the text in depth. We explore and discuss the main topics grouped by thematic interdependence.

**Topic A: Use of big data, demand forecasting & customer satisfaction**

The interaction between man and machine has made it possible to generate ever greater amounts of data which are one of the most valuable assets in tourism (Buhalis and Leung, 2018). The tourism
industry is among those contributing to the generation and use of data flows to the extent that policymakers, tourism destinations, businesses and consumers making up the industry ecosystem are increasingly creating and using BD to improve decision-making processes and co-create value (Wu et al., 2017; Buhalis and Sinatra, 2019; Mariani, 2020). Nevertheless the prominence of the topic of BD in tourism, research on BD in tourism is still rather fragmented and predominantly empirical.

BD is generated through several sources including Internet, mobile transactions, social media, user-generated content as well as purposefully captured content (Xiang et al., 2015). Moreover, the so-called BD analytics approach emphasizes and leverages the capacity to collect and analyse data with an unprecedented breadth, depth, and scale to solve real-life problems (Xiang et al., 2015).

The definition of BD is linked to their ability to describe and/or estimate social, economic and political phenomena (Yang et al., 2021). Many scholars are increasingly interested in adopting BD to better understand complex tourism phenomena, also concerning shocks such as the Covid-19 pandemic. In fact, BD increasingly contributes to the development of the tourism industry in several ways. They offer the opportunity to enrich research designs and methods in tourism researches as they might be progressively embedded also in mixed methods (Miah et al., 2017; Alaei et al., 2019). In relation to that, some studies confirm that the use of BD as a methodology can improve the accuracy of estimates compared to other methods such as questionnaires and interviews, which can also be more costly in terms of time and money (Alaei et al., 2019). Compared to established techniques, they allow a potentially infinite amount of information to be analysed, guaranteeing greater comparability of results. Thus, BD improve the generalizability of research findings across different institutional, economic, social and geographical contexts.

Moreover, the tourism industry, which is a service industry, heavily relies on customer feedback and impressions (Xiang et al., 2015; Cheng et al., 2019; Sun et al., 2019), then measuring satisfaction becomes critically important to understand the degree of customer satisfaction with the services provided, generating also relevant managerial insights and business intelligence (Li et al. 2017; Law et al., 2019; Zhang et al., 2021). Advanced AI applications increasingly need to be fed by BD to trigger learning processes that machines can use to learn from customer behaviours to anticipate and identify future tourists’ needs and engage with travellers (Liu et al., 2017).

In particular, as Buhalis and Leung (2018) assert, the contemporary management of tourist flows requires two types of BD: internal data such as hotel and guest booking history for smart hospitality; and external contextual information such as economic, political and environmental data and profiles of neighbouring events to conduct a comprehensive business analysis. The strategic combination, elaboration and management of these kinds of data allow the creation of a smart and connected tourism environment operating in real time (Buhalis and Sinatra, 2019).
We are just at the beginning of the research and application of BD, and many scholars agree that the research and application of BD in the industry is still at the beginning and it is necessary to improve qualitative and quantitative researches both to understand how BD can further contribute to the advancement of the industry and develop reading and processing methodologies to provide increasingly accurate, comparable and verifiable estimates (Wu et al., 2017; Alei et al., 2019; Galego and Font, 2021).

**Topic B: Augmented and Virtual Reality experience & Value co-creation**

In the 80s, Jaron Lanier firstly coined the expression “virtual reality” (VR), an immersive synthetic 3D simulated environment that allows consumers to have the feeling of being in a real-world environment (telepresence) (Vince, 2004; Gutierrez et al., 2008; Guttentag, 2010; Loureiro et al, 2020). The 1990’s witnessed the development of VR over the Internet and the use of equipment mainly devoted to gamification. Nevertheless the books *Introduction to virtual reality* (Vince, 2004) and *Stepping into virtual reality* (Gutierrez et al., 2008) are considered to be the pillars of the modern conception of VR and AR, in the field of tourism only comes to the fore with the article of Guttentag (2010) *Virtual reality: Applications and implications for tourism*, recognised by many authors to be the pillar in the field (Loureiro et al, 2020; Tussyadiah et al., 2018).

VR is widely defined by Guttentag (2010) as “the use of a computer-generated 3D environment – called a ‘virtual environment’ (VE) – that one can navigate and possibly interact with, resulting in real-time simulation of one or more of the user's five senses.” (p.638). While AR is defined as a type of VR.

Of the main elements constituting the VR environment, the *level of immersion*, is the factor that can most influence the user's *feelings of presence*, which are on the one hand subjective and naturally associated with the user's psychology and, on the other, by the ability of AR systems to deliver high quality (Guttentag, 2010; Tussyadiah et al., 2018). Thus, it links a strong subjective component of the user in receiving and high objective quality of data and tools to simulate through sensors such as input devices that interpret the user's actions and can respond contextually. A wide variety of VE can be created and found, including online, of which, one of the most widely used is Second Life where user-avatars can explore VE through social interactions, entertainment and commerce (Guttentag, 2010; Tussyadiah et al., 2018).

If in the field of tourism, the first application was found in the project Rome Reborn, a 3D virtual model of ancient Rome, which uses virtual sites’ reconstructions and 3D edutainment through avatars (Guttentag, 2010). Then, several other applications of VR in tourism areas have been carried out with
extraordinary results. The first area of application of AR in tourism is in planning and management, where it can have several implications. In particular, we highlight the power of VR to aid in policy formulation through the power of being able to virtually observe an environment and its changes. Furthermore, it allows for the communication of plans to the members of a community through direct immersion (Guttentag, 2010). Secondly, marketing is one of those that can benefit the most, due to its high capacity to sponsor a destination, tourism products or services by attracting tourists and consumers 'remotely' (Guttentag, 2010; Tussyadiah et al., 2018; Kim et al., 2020; Bogicevic et al., 2020). Examples of this are virtual tours that allow immersion in the synthetic reality of the tourism product. Moreover, VR offer unique platforms to communicate among tourists with a higher degree of trust (Buhalis and Law, 2008; Buhalis and Sinarta, 2019). The third area of application concerns entertainment. The field of gaming is one of the first to emerge in connection with VR themes (Sensorama Simulator, 1962) and one of the most promising. Although many entertainment applications in VR are designed for home use, others have already established themselves as tourist attractions in tourist destinations. One example is theme parks, where the application of AR tools allows to experience environments and worlds far away in time and space (Jung et al., 2015; Wei et al., 2019). Linked to this and fourth field of application is education, as AR is a powerful interactive tool for teaching and learning directly from the 'synthetic' experience (Guttentag, 2010; Kim et al., 2020). Finally, several researchers have suggested that VR could serve to preserve cultural heritage by providing an alternative form of access to fragile sites and environments (Guttentag, 2010; Tussyadiah et al., 2018). Firstly, the rendering of sites through virtual 3D models can be a valuable tool for cultural heritage conservation, as these virtual models can contain extremely precise and accurate data sets that, in theory, can be preserved indefinitely, while a site or object may be degraded by impacts such as erosion or deterioration. Furthermore, VR can protect sites, such as UNESCO World Heritage sites that attract large numbers of visitors, from deterioration caused by over tourism issues in fragile environments while preserving the sustainability of places (Guttentag, 2010; Tussyadiah et al., 2018).

Not only VR and related technologies are making a revolution in the way tourists experience travel and tourism-related products and services, but also may be used as powerful tools to co-create value with users challenging the tourism industry itself (Dieck and Jung, 2018; Tung and Au, 2018; Buhalis and Sinarta, 2019). Co-creation of value is a theme related to the direct immersion of the consumer in the tourist experience and the possibility of interaction between producer and user. Buhalis and Sinarta (2019) identify how brands in tourism utilize digital technology to enhance consumer experience in real-time to co-create value, i.e. “joint creation of value by the company and the customer” (Prahalad & Ramaswamy, 2004a, p.8). Co-creation is about creating an experience
environment where each member of the ecosystem is highly involved in the value creation and thus, enables consumers to co-construct their own unique personalised experience (Buhalis et al., 2019). ICTs such as VR and AR tools may act as a key medium to engage in real time with consumers (Buhalis and Foerste, 2015; Buhalis and Sinarta, 2019).

**Topic C: Covid-19 pandemic, health care and social distances & service robots**

In the field of tourism, an industry that bases its strength on the ability and possibility of visitors to experience faraway places and that requires a high level of human interaction, the Covid-19 pandemic was the cause of a total shutdown (Zeng et al., 2020). Since 2020, numerous studies, both empirical and theoretical, have been conducted to address the threats and challenges posed by this event. Although the authors agree that it is still too early to know exactly what travel and tourism will be like after we finally emerge from isolation, scholars agree that ICT employed to manage the disease will continue to play a significant role (Jiang and Wen, 2020; Zeng et al., 2020).

New AI forecasting methodologies (Fotiadis et al., 2021) and technologies have led the industry in solving some of the problems related to the pandemic in various strategic sectors such as transport, hospitality, catering and access to tourist sites and attractions. In this regard, it emerges that a conspicuous strand of research on AI and tourism concerns the robotization of the industry in terms of consumer preference and attitude to interface with human-like robot (de Kervenoael et al., 2020; Zeng et al., 2020; Pillai et al., 2021).

Robots have often been used in response to disasters and crises, such as the Fukushima nuclear power plant meltdown (Zeng et al., 2020). However, with advances in AI and other technologies, robotics has become increasingly viable in tourism to provide concierge, cleaning, catering and other services (Cain Lisa et al., 2019; Ivanov et al., 2019; Jiang and Wen, 2020; Yu, 2020).

In the COVID-19 pandemic, robots initially proved particularly effective for routine or non-routine tasks, such as using ultraviolet light to disinfect surfaces. Later, however, robot technology matured rapidly, and different types of robots appeared to handle COVID-19 in various environments, including hospitals, airports, transport, recreational and scenic areas, hotels and communities in general. In particular, robots have taken on some essential tasks in protecting the community through disinfecting and monitoring spaces, ensuring social distancing imposed through autonomous transport and delivering necessities, within hospitals, airports and stations to assist managers in ensuring safety and distances, provide real-time information and directional services.

The technological drift of the industry brings contradictions and consequences (Lu et al., 2019). Seyitoglu and Ivanov (2020), assert that empathy and emotional intelligence are important drivers of
customer perceptions, service experience and quality. However the use of service robots may increase the psychological distance between tourists and residents in a destination due to the decreased physical interaction between them. Kim et al. (2021), applying four experiments to research tourists’ liking for robots over humans, showed that the pandemic shifted liking towards services provided by robots. Tung et al. (2018) suggest that human-machine interaction can lead to the co-creation of values and it is therefore important to assess areas where the application of machines can positively influence consumer perceptions. Murphy et al. (2019) by addressing the anthropomorphism of machines suggest intervening directly on those features that anthropomorphize machines to decrease the distance with customers.

The second argument about service robots concerns their convenience and efficiency compared to humans. Indeed, service robots can reduce operating costs and provide customers with better services. Belanche et al. (2021) found that the customer's affinity with the service robot positively influences the attribution of service improvement, which in turn has a positive influence on the customer's behavioural intentions. Conversely, affinity has a negative influence on the attribution of cost reduction, which in turn has a negative effect on behavioural intentions. Finally, human similarity has a positive influence on affinity.

We are certainly seeing a shift in terms of perception between pre- and post-pandemic. The industry should push towards robotisation in those phases where user interaction is low or non-existent such as cleaning services and focus on human interaction in those services where face-to-face dialogue can enhance the experience itself such as catering and hospitality (Seyitoglu and Ivanov, 2020).

**Topic D: Smart tourism trends**

In recent years, a new and global search stream has emerged under the umbrella of the smart tag which has found one of the greatest strength in conjunction with the context of cities to describe an environment where city elements and technologies interact in innovative ways to achieve resource optimization, effective and equitable governance, sustainability and quality of life: smart city (Gretzel et al., 2015; Adler and Florida, 2018). What remains unclear is the definition of the smartness concept which has been analysed in many disciplines. Smartness is defined as a new buzzword to describe technological, economic and social developments supported by technologies that rely on sensors, BD, open data, new modes of connectivity and information (e.g. Internet of Things, Internet of Everything) and the ability to infer and reason through complex methodologies such as deep learning, machine learning, and so on. (Gretzel et al.; 2015). There is a lack of clarity around smart tourism
definition: “suddenly everything is smart” (Gretzel et al.; 2015, p.180). If *smart* is therefore used to describe a complex mixture of several interconnected elements, certainly, the technological part holds the central position. Both Buhalis and Leung (2018) and Adler and Florida (2018) state that the concept of smartness refers to a network of (public and private) organisations and smart functions that engage in interoperable and interconnected systems to simplify and automate everyday activities and (co-) create value across the ecosystem for all stakeholders.

In the tourism industry the term *smart tourism*, but also *smart destination* or *smart tourism experience*, is becoming more and more fashionable, as shown by the size of the word *smart* in the keyword co-occurrence analysis (section 4.3). Many of the smart tourism initiatives have arisen from smart city projects and, consequently, from smart tourist destinations (Gretzel et al., 2015; Adler and Florida, 2018).

Gretzel et al. (2015) identify three main layer of smart tourism: *smart destinations* which are special cases of smart cities that considered not only residents but also tourists in manage mobility, resource availability and allocation, sustainability and quality of life/visits; *smart experience* that focuses particularly on technology-mediated tourism experiences and their enhancement through personalisation, context awareness and real-time monitoring, and where users contribute data through photos, feedback, etc.; finally, the third layer is the *smart business* which refers to the complex business ecosystem that creates and supports the exchange of touristic resources and the co-creation of the tourism experience. Adler and Florida (2018), focus their study right on smart tourism experience which should include smart technologies and solutions to offer data-driven, built in real time, context-aware and co-created experiences.

Buhalis and Leung (2018) focus on the smart hospitality ecosystem, which involve a large number of direct and indirect stakeholders, proposing an innovative framework at the heart of which two main elements are identified: interoperability and interconnectivity. Indeed, smartness can effectively develop networks to create an ecosystem and dynamically interconnect all members. Interoperability refers to the ability of business systems and processes to support the exchange of data and enable the sharing of information and knowledge. Interconnectivity means that the barriers to collaborate are minimized, effectively assisting hotels to constantly evaluate which subsystem is serving their needs and strategies better. Although the focus is on hospitality, the authors state that management can easily apply technology to support tourism “*as a highly-interconnected and networked industry***” (Buhalis and Leung, 2018, p.42) and empower the co-creation process.

One of the main problems identified in smart tourism is the lack of (big) data organization and standardization and data access. Furthermore, Gretzel et al. (2015) state that tourism is still not a
sector that attracts many knowledge workers and despite its heavy reliance on ICT, it is known to have problems with a lack of innovation.

The development of smart tourism is certainly already underway due to the widespread deployment of technology in the industry. However, systematic coordination and the exploitation of data sources such as BD for value creation are still in their infancy. Although smart tourism initiatives around the world are promoting the construction of viable smart tourism ecosystems (Gretzel et al. 2015), the complexity of the industry makes it very difficult to go beyond specific innovations. Further research needs to be conducted for clarity about the meaning of the term, to define the elements that make up the ecosystem for smart tourism, and quantitative studies need to be conducted to explore the ecosystem through new statistical methodologies.

5. Conclusions

The present research aimed to investigate the development of AI in scientific research in tourism. The work developed a joint bibliometric analysis with an analysis of keywords co-occurrence and network analysis of the main authors, along with a qualitative analysis of the literature to approach the topic from different perspectives.

While there have been a couple of wide literature reviews covering BD in hospitality and tourism up to 2017 (Li et al., 2018; Mariani et al., 2018), to the best of our knowledge there is only one study on AI in tourism research (Lv et al., 2021). This study thus provides further insights into the evolution on research and presents an up-to-date overview.

Our first result confirms the research of Mariani (2020) and Lv et al. (2021) about the exponential growth of studies on AI and tourism in recent years. However, our interdisciplinary focus has allowed us to point out that the journals involved in the debate are not only specialists on tourism topics, and that tourism has undergone a disciplinary broadening in this regard (Mariani and Baggio, 2021). However, in recent years this phenomenon is changing and AI and tourism are also registering an increasing centrality in tourism journals (Tourism Management is the first publishing journal on the topic). Like other authors, we have highlighted the presence of many contributions that use AI and BD exclusively as tools for tourism research. However, exclusive fields in which AI and tourism can be developed were also highlighted: indeed, we discussed the opportunities and challenges that COVID19 has created and how they can be addressed through AI applications such as AR, VR and robots.

The interdisciplinarity of the topic is also confirmed by the network of authors. Indeed, although some networks of collaborating authors have been identified, the study shows that networks of collaborations are still not particularly intense, perhaps also due to the youth of the topic itself.
An interesting result concerns the topics most discussed in the scientific literature which would seem to constitute the directions in which the tourism industry is moving. Among these, four main thematic blocks have been identified: the first revolves around the role of BD for demand prediction and consumer care; the second concerns the use of new tools aimed at improving the consumer experience, AR and VR; the third group looks at the recent Covid-19 pandemic and the possibilities for the industry, particularly in the field of robotics; finally, the fourth concerns the theme of smartness applied to the tourism industry with a perspective of developing an increasingly interconnected industry.

In particular, the use of AI-based tools has advanced tourism research, enabling the use of BD, which is particularly useful in demand forecasting, online customer satisfaction, and all studies of social networks and online data. This poses challenges, however, in businesses and tourism organizations and institutions for the integration of business intelligence (Mariani, 2020) and to develop expertise in these tools.

The use of AI in tourism offers many opportunities, but it also needs to be accompanied by better training of staff in universities, in developing teams with more analytical and multidisciplinary skills otherwise struggling with new trends in industry development and data analysis.

It is not all positive however, in this context there is also a dark side of AI and this is also the case of tourism research. Lazzeretti (2021) highlights the problems related to the nature of the algorithms and the black box principles of their operation and, as well as the use of data in the wild and biases and hidden errors, etc. The use of data created for non-specific purposes, the use of non-transparent tools (random forest or neural network), hinder an even wider dissemination of these tools.

Certainly, the tourism industry can also benefit greatly from new technologies and the use of AI and BD as a service industry, very much oriented to online platforms, tourist reviews, customer satisfaction, and the need to forecast the demand and predict consumer behaviour. However, like other disciplines, it faces some challenges and dangers that will need to be addressed.

In conclusion, this article has provided an up-to-date overview of AI research in scientific research in tourism, describing the main trends of study, the most important authors, the most used journals and the most developed techniques. We are at a very relevant historical moment, AI and BD methodologies are no longer borrowed from other fields and applied occasionally to tourism but are increasingly developed and central to published works and studies. The future is expected to confirm this trend, and the discipline and scholars will have to equip themselves with skills and capabilities that are increasingly needed in a society that is constantly changing technologically, and tourism cannot stand aside.
References


