UNIVERSITY OF LJUBLJANA
SCHOOL OF ECONOMICS AND BUSINESS

MASTER'S THESIS

ELLA POMMERANZ

EUROPEAN MASTER IN TOURISM MANAGEMENT

FACULTY OF HUMANITIES, UNIVERSITY OF SOUTHERN DENMARK
SCHOOL OF ECONOMICS AND BUSINESS, UNIVERSITY OF LJUBLJANA
FACULTY OF TOURISM, UNIVERSITY OF GIRONA

MASTER'S THESIS

OPPORTUNITIES AND CHALLENGES OF LEVERAGING BIG DATA IN SMART TOURISM DESTINATIONS

Ljubljana, July 2022 ELLA POMMERANZ

AUTHORSHIP STATEMENT

The undersigned Ella Jana Pommeranz, a student at the University of Ljubljana, Faculty of Economics, (hereafter: FELU), author of this written final work of studies with the title "Opportunities and Challenges of Leveraging Big Data in Smart Tourism Destinations", prepared under supervision of Professor Ljubica Knežević Cvelbar.

D E C L A R E

1. this written final work of studies to be based on the results of my own research;

2. the printed form of this written final work of studies to be identical to its electronic form;

3. the text of this written final work of studies to be language-edited and technically in adherence with the FELU's Technical Guidelines for Written Works, which means that I cited and / or quoted works and opinions of other authors in this written final work of studies in accordance with the FELU's Technical Guidelines for Written Works;

4. to be aware of the fact that plagiarism (in written or graphical form) is a criminal offence and can be prosecuted in accordance with the Criminal Code of the Republic of Slovenia;

5. to be aware of the consequences a proven plagiarism charge based on the this written final work could have for my status at the FELU in accordance with the relevant FELU Rules;

6. to have obtained all the necessary permits to use the data and works of other authors which are (in written or graphical form) referred to in this written final work of studies and to have clearly marked them;

7. to have acted in accordance with ethical principles during the preparation of this written final work of studies and to have, where necessary, obtained permission of the Ethics Committee;

8. my consent to use the electronic form of this written final work of studies for the detection of content similarity with other written works, using similarity detection software that is connected with the FELU Study Information System;

9. to transfer to the University of Ljubljana free of charge, non-exclusively, geographically and time-wise unlimited the right of saving this written final work of studies in the electronic form, the right of its reproduction, as well as the right of making this written final work of studies available to the public on the World Wide Web via the Repository of the University of Ljubljana;

10. my consent to publication of my personal data that are included in this written final work of studies and in this declaration, when this written final work of studies is published.

Ljubljana, July 15th, 2022 Author's signature: E. Pommeranz

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AI – Artificial Intelligence

API – Application Programming Interface

AR – Augmented Reality

BDA – Big Data analysis

COVID-19 – Coronavirus disease 2019

EU – European Union

GDP – Gross Domestic Product

GPS – Geographical Positioning Systems

HOSBEC – Hotel Business Association of Benidorm, Costa Blanca and the Valencian Community

ICT – Information and Communication Technology

INE – Instituto Nacional de Estadística (National Statistics Institute of Spain)

INVATTUR – Valencian Institute of Tourist Technologies

IoT – Internet of Things

IT – Information Technology

KPI – Key Performance Indicator

NFC – Near Field Communication

PMS – Property Management System

RFID – Radio-frequency identification

SEGITTUR – Sociedad Mercantil Estatal para la Gestión de la Innovación y las Tecnologías Turísticas (State Trading Company for the Management of Innovation and Tourism Technologies)

SIT – Tourism Intelligence System

SME – Small- and medium-sized enterprise

SSI – Semi-structured interview

STD – Smart Tourism Destination

U4SSC – United for Smart Sustainable Cities

UGC – User-generated content

UNWTO – United Nations World Tourism Organization

VR – Virtual Reality

WTTC – World Travel & Tourism Council

ABSTRACT

Advancements in Information and Communication Technology (ICT) in recent years have altered practically every industry, including the tourism sector, which has driven the creation of Smart Tourism Destinations (STDs) that are founded upon such technological developments. These also led to a growing volume of data being produced, from which the concept of Big Data originated. However, these two novel notions have yet to be linked further within research. By conducting seven in-depth interviews with tourism professionals from the private and public sectors in the STD of Valencia, Spain, this research examines the potential and obstacles that STDs may face when employing Big Data. The findings indicated that Big Data enables better forecasts and decision-making, which may aid in destination and business management and improve customer experiences. However, before attaining those benefits, companies must overcome significant obstacles at practically every stage of the data analysis process, including data collection, sharing, integration, and data security. Although there is no universal approach, the results suggest that destinations and businesses must first fully embrace digitalization and acquire the financial, human, infrastructural, and organizational enablers that are essential for utilizing this new technology. While data collection and infrastructure may be outsourced, it is recommended that tourism actors progressively drive internal change through education regarding basic data analysis capabilities. This will be crucial to derive insights and value from the obtained data, thereby increasing the understanding of and minimizing the reluctance to work with Big Data.

Keywords: Big Data, Smart Tourism Destinations, Valencia, ICT

INTRODUCTION

At the start of 2020, the amount of data bytes in the digital space was 40 times greater than the number of stars in the visible universe (Desjardins, 2019). Due to fast-paced technological advancements such as the introduction of the internet, more and more data is continuously being generated and stored (Kambatla et al., 2014; Katal et al., 2013; J. Li et al., 2018; Verhoef et al., 2016). Furthermore, the way information is created and consumed has evolved dramatically, including in the tourism industry (Alaei et al., 2019). A multitude of tourism-related data is constantly being created, such as information on lodging, points of interest, and reviews (Figueredo et al., 2018), resulting in a plethora of options and thus decisions for travelers and travel businesses alike (Kontogianni & Alepis, 2020). Yet, manual processing of this large volume of data became practically impossible, necessitating new analytical and processing techniques, which ultimately led to the emergence of the Big Data concept (Alaei et al., 2019; M. Chen et al., 2014; Mariani et al., 2018).

Prior to the outbreak of the Coronavirus disease 2019 (COVID-19) pandemic, which has caused an unforeseen recession in global economic activity bringing the entire world to a halt (Khalid et al., 2021), tourism was one of the fastest-growing sectors worldwide, according to the World Travel & Tourism Council (WTTC, 2019a). It is a significant contributor to economic growth (UNWTO & UNDP, 2017) and accounted for one in every ten jobs worldwide as well as contributed to ten percent of the global Gross Domestic Product (GDP) before the pandemic (WTTC, 2019a, 2020). Nevertheless, recently there has been a shift to more sustainable tourism development as a return to the status quo pre-pandemic, i.e., indefinite growth, has proven to be unsustainable (Liburd & Edwards, 2018). Further, the pandemic led to an increased demand for digital solutions that allow for the monitoring of customer and resident flows (Gretzel & Koo, 2021), the generation of insights and the prediction of demand (Sigala et al., 2012; Song & Liu, 2017). One means of attaining these objectives is Big Data analysis (BDA).

The analysis of massive amounts of data is a valuable tool to understand the consumer market better, personalize offers, and improve the tourism experience, and may contribute to improving destination smartness through “smarter” decisions in areas such as destination planning, strategy, and image (Del Vecchio et al., 2018; Gajdošík, 2019; Marine-Roig & Anton Clavé, 2015; Verhoef et al., 2016; Xiang, Tussyadiah, et al., 2015; Xiang & Fesenmaier, 2017b). This enormous and diversified data should be made operational for destinations and taken into consideration when making future development choices as it may help smart cities turn into Smart Tourism Destinations (STDs) (Marine-Roig & Anton Clavé, 2015). After the recent aim for sustainability, destinations are expected to next transform into smart destinations by focusing on technology and innovation (Dexeus, 2021). The notion of being “smart” has been a buzzword in recent years and is now being applied to tourism development as well (Gretzel, Sigala, et al., 2015; Ivars-Baidal et al., 2019; Sigalat-Signes et al., 2020; W. Zhu et al., 2014). It outlines how the respective industries, tourism destinations, and their visitors are becoming progressively more reliant on new types of

Information and Communication Technologies (ICTs) (Gretzel, Sigala, et al., 2015). These allow for big volumes of data to be collected, aggregated, and finally turned into value propositions "with a clear focus on efficiency, sustainability and experience enrichment" (Gretzel, Sigala, et al., 2015, p. 181).

Despite the recognition of the benefits of Big Data in the tourism industry, research is limited, both theoretically and practically, regarding the adoption of BDA faced by STDs (Ardito et al., 2019; Gretzel, Sigala, et al., 2015; Kontogianni & Alepis, 2020; Song & Liu, 2017) as the majority of literature centers around smart cities (Buhalis & Amaranggana, 2013). So far, only few empirical studies have been undertaken to evaluate the true capabilities of Big Data (Fosso Wamba et al., 2015), and the term, its potential, and commercial value remain unclear to some (Fosso Wamba et al., 2015; Gandomi & Haider, 2015; Mariani et al., 2018; Sheng et al., 2017). Further, the systematic collection, integration, and exchange of touristic data are still in the early stages, and the sector's complexity makes it challenging to move beyond platform-, technology- or service-specific advancements (Gretzel, Sigala, et al., 2015). That is why in current research, data analytics and STDs have yet to be linked as well as be looked upon from a more critical stance as past research tended to be utopian (Gretzel, Reino, et al., 2015; Gretzel, Sigala, et al., 2015). The use of data within STDs necessitates further exploration and conceptualization from a variety of disciplines and research methods (Del Chiappa & Baggio, 2015; Gretzel, Sigala, et al., 2015) both to advance theoretical contributions and to validate prior studies (Buhalis & Amaranggana, 2013). Considering that Big Data is a new technology that has the potential to provide significant operational and strategic benefits to businesses, it is vital to also raise awareness of the different obstacles slowing down the adoption of this technology (Katal et al., 2013). Following this rationale, the research question for this thesis follows as:

"What opportunities and challenges are Smart Tourism Destinations facing when leveraging Big Data?"

This research attempts to fill the knowledge gap in the literature by building on previous research as well as an in-depth case study. By examining the implementation of Big Data in the smart destination of Valencia, Spain, this research investigates the potential obstacles and opportunities of smart destinations using Big Data. This research is carried out in a Western context from a managerial perspective of tourism professionals. The city of Valencia was chosen as a case study as it has been awarded the European Capital of Smart Tourism 2022 and is said to have developed a holy grail in terms of an open tourism data platform (European Commission, n.d.-a). This could provide helpful insights and act as a guideline for preparing tourism destinations in a similar context aspiring to become "smarter" in the future through the use of Big Data. After reading this thesis, the reader will have a clearer understanding of the role, promises and related challenges of Big Data in a tourism context.

Further objectives of this thesis are to:

- clarify the definition and concepts related to Big Data and smart destinations
- summarize existing literature on how Big Data may be utilized in a tourism context
- conduct an exhaustive analysis of the case study destination of Valencia, Spain by analyzing the opinions of tourism professionals through in-depth interviews
- determine the challenges and opportunities of Big Data
- identify key aspects to consider when employing Big Data in a smart destination
- contribute to the scarce tourism literature on this subject

The research problem is investigated using an exploratory research methodology, which is the preferred method for studying phenomena with little data (Jennings, 2015; Mason et al., 2010). Primary data from seven semi-structured in-depth interviews with tourism professionals from Valencia's public and private sectors is used to answer the research question and is being examined using thematic content analysis to discover common themes.

Following this introduction, the first chapter of this study will delve into the theoretical foundation of this topic by defining and linking key concepts such as Big Data and Smart Tourism Destinations, as well as depicting their evolution over time. The employed theoretical paradigm and approaches will be explained within the methodology chapter (2). This is followed by a presentation of the case study of Valencia, Spain, which describes the city's transformation into a smart city and tourism destination. Chapter 4 examines the data acquired through interviews and explains the findings. At this point, the results will be discussed in relation to what has been stated in the literature presented in the first chapter. Lastly, the thesis will be concluded by addressing the research question, highlighting the study's contributions, shortcomings, and recommendations for future research.

1 THEORETICAL BACKGROUND

The aim of this chapter is to produce a literature review centered around the research question in order to lay a theoretical framework for the thesis. It is a necessary component of empirical investigations (Alvesson & Sköldberg, 2000) and outlines the research by defining key terminology and concepts that will be covered throughout the dissertation. Smart Tourism Destinations and Big Data are two important terms whose definitions and relationships will be explored in the following sections. First, the published research on STDs and Big Data is reviewed. Next to each concept's definition, the evolution of the two notions is also described. This allows the reader to comprehend the concepts and learn about their possibilities as well as drawbacks, both in general and in the tourism context.

1.1 Smart Tourism Destinations

1.1.1 Evolution of the tourism destination concept

Tourism is "a social, cultural and economic phenomenon which entails the movement of people to countries or places outside their usual environment for personal or business/professional purposes" (UNWTO, n.d.-b, sec. 1). These people visit so-called **tourism destinations** which are defined by the United Nations World Tourism Organization (UNWTO) as "[...] a physical space with or without administrative and/or analytical boundaries in which a visitor can spend an overnight. It is the cluster (co-location) of products and services, and of activities and experiences along the tourism value chain [...]. A destination incorporates various stakeholders […]" (UNWTO, 2019, p. 14). The concept of tourism destinations has developed over time (Jovicic, 2019). During the early development stages of destinations, the initial tourism infrastructure, products and services are established (Dexeus, 2021; cf. Figure 1). After, destinations begin with strategic planning and then move to focus on sustainability as well as the improvement of the quality of services as destinations mature. The most recent advancements in tourism development models are posed by smart destinations which were created upon the accumulated knowledge and experience of the past and attempt to tackle contemporary challenges. They center around the aspects of technology, innovation and knowledge, which require investment in Information Technology (IT) infrastructure.

Figure 1: Evolution of the tourism destination concept

Source: Dexeus (2021, p. 3).

Framke (2002) and Jovicic (2019) also explored the evolution of tourism destination concepts from the 1970s onwards and mainly distinguished between the traditional business-oriented concept of tourism destinations and the systemic approach to tourism destinations, which includes a sociological understanding. Jovicic (2019), however, expanded this research and additionally included the concept of the smart destination. The classical perspective on destinations emphasizes the importance of geographical factors such as attractions, accommodation and transportation infrastructure (Framke, 2002; Georgulas,

1970; Howie, 2003; Jovicic, 2019). This is similar to what Buhalis (2000) referred to when he introduced six elements that must be present in order for a tourism destination to be successful, which are also known as the 6 As of tourism destinations. These can be summarized as attractions, accessibility (i.e., transportation infrastructure), amenities (i.e., all services allowing a comfortable stay, such as lodging or food), activities, available packages (i.e., service bundles provided by intermediaries) and lastly, ancillary services (i.e., services such as banking, postal service, and hospitals that are not primarily aimed at tourists). In the traditional view of tourism destinations, destinations were thus mostly seen as spatial units or clusters of attractions and services (Burkart & Medlik, 1981; Mill & Morrison, 2012) whilst disregarding collaboration among stakeholders within the destinations as well as the active role of tourists (Framke, 2002; Jovicic, 2019). A major shift in the conventional view of destinations happened in the mid-1990s as tourism practice and theory advanced and tourism researchers advocated a systemic and more holistic approach to tourism (Jovicic, 2019). Next to the connection of the tourism industry with the macro-environment (Howie, 2003; Wall, 1996), the aspect of interaction between tourists, tourism companies, and residents became crucial. Tourists were then being paid more attention to and recognized as experience-seeking individuals whose subjective perceptions of and demands for a destination matter in destination management (Buhalis, 2000; Framke, 2002).

During the same time period, the concept of **smartness** was born. It was first mentioned in the 1990s and quickly became a buzzword after 2008, describing the technical, economic, and social changes caused by the interconnection of sensors, Big Data, and other technologies such as the Internet of Things (IoT) and Near Field Communication (NFC) (Gretzel, Sigala, et al., 2015; R. Hollands, 2014). Smartness, according to Harrison et al. (2010), is defined as leveraging useful, near real-time data and making better operational decisions by integrating and exchanging data, using advanced analytics, optimization, and visualization. The widespread use of digital technologies has drastically altered many parts of life and had a massive influence on various economic sectors (Baggio et al., 2020). These ICTs add value to their users as every actor can easily access knowledge and information through their utilization (Zhou & Li, 2012). The concept of "intelligence" or "smartness" has since then been ascribed to many other areas such as physical infrastructure (e.g., smart homes), markets (e.g., smart economy), technologies (e.g., smartphone), tourism or cities, describing the enhancement of these fields through the incorporation of technology (Boes, Buhalis, et al., 2015; Gretzel, Sigala, et al., 2015).

Smart cities, for example, have been investigated since the 1960s (Angelidou, 2014). A smart city is a multidisciplinary concept that combines technology, innovation, information, governance, economy, society, infrastructure, education, sustainability, and quality of life (Avelar, 2020). It thereby also builds the basis for a Smart Tourism Destination (ibid.) **Smart tourism**, and consequently STDs, as a concept emerged much later after IBM introduced the concept of a 'Smarter Planet' in 2008 (IBM, 2015) that centered around cloud computing, mobile technology, Big Data, and analytics to transform businesses. As the concepts of smart

cities and STDs are closely interlinked and often used interchangeably, the two phenomena will be defined more precisely in the subsequent chapters.

1.1.2 Smart cities

One of the twenty-first century's key problems is to construct sustainable and interconnected cities (Sigalat-Signes et al., 2020), which is why many global metropolitan cities have begun to concentrate on sustainable growth through the regulation of urban planning (Trencher & Karvonen, 2019). Urban areas have expanded swiftly (Gretzel & Koo, 2021; Koo et al., 2021) and 60% of the world's population is expected to be living in megacities or dense metropolitan regions by 2030 (Yigitcanlar & Kamruzzaman, 2018). Yet, this urban sprawl has come at a cost, including increased pressure on infrastructure, higher complexities, and social and environmental concerns (Falconer & Mitchell, 2012; Gretzel & Koo, 2021; Sigalat-Signes et al., 2020). While the smart city concept has already existed since the 1960s (Angelidou, 2014), the popularity of the concept skyrocketed in international legislation and academic studies within the last two decades as a potential solution to the issue of sustainability and as an answer to the major problems of vast urbanization (Achaerandio et al., 2011; Albino et al., 2015; Boes, Buhalis, et al., 2015; Buhalis & Amaranggana, 2013; Gretzel & Koo, 2021; Hall et al., 2000; Jasrotia & Gangotia, 2018). Thus, the smart city framework has been regarded as the new ideal for urban development in recent years (Sigalat-Signes et al., 2020) leading to several cities throughout the world taking steps to establish themselves as smart cities (Gretzel et al., 2016; Khomsi & Bedard, 2016).

But what exactly defines a smart city? There is no unified global definition of a "smart city" since the existing ones are ambiguous and inconsistent, resulting in a discrepancy in meanings that vary depending on the perspective (Al Nuaimi et al., 2015; Alastal & Badawy, 2020; Avelar, 2020; Boes et al., 2016; Meijer & Rodríguez Bolívar, 2015; Nam & Pardo, 2011; Sigalat-Signes et al., 2020). One could even say that the understanding of the smart city idea differs not only by expertise and point of view but also from city to city (Alastal & Badawy, 2020). Nevertheless, the definition of Caragliu et al. (2011) is one of the most widely recognized by the research world, according to Sigalat-Signes et al. (2020). These researchers argue that a city is considered to be smart when investment in human capital, involvement in governance, and infrastructure that supports the information distribution throughout the city are being used to create long-term economic growth and high quality of life. The smart city's objective is to encourage sustainable socio-economic growth, the improvement of urban services like transportation and the effective use of resources via the adoption of ICTs (Alastal & Badawy, 2020; Cimbaljević et al., 2019; European Commission, n.d.-b; Gretzel, Sigala, et al., 2015; Sigalat-Signes et al., 2020). A smart city is constituted of and monitored by ubiquitous ICTs (Neirotti et al., 2014), which is why ICTs are frequently regarded as the key component of a smart city (Nam & Pardo, 2011). These ICTs bring all services together, making citizens more connected, informed, and involved (Buhalis &

Amaranggana, 2013). The technologies are also utilized to make cities both more accessible and pleasant for locals as well as visitors by integrating all local actors to provide real-time solutions and use data collectively for improved coordination (Buhalis & Amaranggana, 2013), hence improving residents' quality of life (Boes, Buhalis, et al., 2015; Gretzel & Koo, 2021). While some authors argue that ICTs might be vital for the functioning of a smart city, others state that they are merely facilitators or tools rather than the foundation for success (Boes, Buhalis, et al., 2015; Caragliu et al., 2011; Khomsi & Bedard, 2016; Nam & Pardo, 2011). Instead, it is stated that sustainability, human capital and an improved quality of life of residents, economic growth, efficient and integrated management, and innovation are core constructs of smartness (cf. Table 1).

Table 1: Core components of smart cities

Component	**Sources**
ICTs	Boes, Buhalis, et al. (2015); Caragliu et al. (2011); Khomsi & Bedard (2016); Nam & Pardo (2011)
Sustainability and smart environment	Al Nuaimi et al. (2015); Gretzel, Sigala, et al. (2015b); Gretzel, Werthner, et al. (2015); Lamsfus et al. (2012); Sigalat-Signes et al. (2020)
Improved quality of life of residents (smart living and smart people)	Al Nuaimi et al. (2015); Boes, Buhalis, et al. (2015); Caragliu et al. (2011); Cohen (2011) as cited in Peek & Troxler (2014); Gretzel, Sigala, et al. (2015a); Gretzel, Werthner, et al. (2015); Nam & Pardo (2011); Sigalat-Signes et al. (2020); Yigitcanlar et al. (2018)
Smart economy	Cohen (2011) as cited in Peek & Troxler (2014); Nam & Pardo (2011)
Smart management and governance	Boes, Buhalis, et al. (2015); Cohen (2011) as cited in Peek & Troxler (2014); Giffinger et al. (2007); Gretzel, Sigala, et al. (2015a); Gretzel, Werthner, et al. (2015); Nam & Pardo (2011); Yigitcanlar et al. (2018)
Innovation	Boes, Buhalis, et al. (2015); Caragliu et al. (2011); Lamsfus et al. (2012); Nam & Pardo (2011); Yigitcanlar et al. (2018)
Smart mobility	Cohen (2011) as cited in Peek & Troxler (2014)

Source: Own work.

Nonetheless, while the concept holds great promises, governments worldwide are worried about the expense of becoming a smart city, as well as the shortage of natural and human resources and skills. Another issue is the regulatory frameworks, which significantly impact the implementation's success (Al Nuaimi et al., 2015).

1.1.3 Smart tourism

In current research, smart tourism is still only vaguely defined as a concept (Gretzel, Sigala, et al., 2015; Xiang, Tussyadiah, et al., 2015). Nevertheless, following Gretzel, Sigala et al.'s (2015, p. 179) definition, smart tourism describes "the increasing reliance of tourism destinations, their industries and their tourists on emerging forms of ICT that allow for massive amounts of data to be transformed into value propositions". Gajdošík (2019) and Gretzel (2011) define smart tourism as the prevailing phase of tourism development impacted by the advancement of digital technologies in tourism, in which the physical and management dimensions of tourism are entering the digital landscape.

Over the past years, smart tourism has gained popularity in both the tourism sector and academia (Gretzel, 2018; Gretzel, Sigala, et al., 2015). According to Google Trends (n.d.), the worldwide interest in searching for the term "smart tourism" has steadily increased in the last five years. While in 2018, solely 2,470 results were shown when looking for the keyword on Google Scholar (Gretzel, 2018), by the end of January 2022, this number already amounted to around 440,000 articles. This trend has also been picked up by numerous governments around the world, which are taking measures to implement smart tourism into their agendas (Gretzel, Sigala, et al., 2015). However, these actions as well as the degree to which smart tourism has been accepted as a vision for tourism differ by region (Gretzel, 2018; Ivars-Baidal et al., 2019). In Asia, for example, governments are investing in developing the needed technology infrastructure for smart tourism (Guo et al., 2014; Hwang et al., 2015; Wang et al., 2013), while in Australia, the focus lies mainly on smart governance and open data (Gretzel, Sigala, et al., 2015). Europe, on the other hand, puts the emphasis on innovation and competitiveness, as well as the enhancement of tourism experiences by processing data in novel ways (Baggio et al., 2020; Boes, Borde, et al., 2015; Boes, Buhalis, et al., 2015; Lamsfus et al., 2015). The European Union (EU) has launched the "European Capital of Smart Tourism" award, which intends to advocate smart tourism in the EU by honoring achievements and boosting destinations through best practice sharing (European Commission, n.d.-a). In the North and South American as well as African contexts, the research on this topic remains very limited, to the best of the author's knowledge.

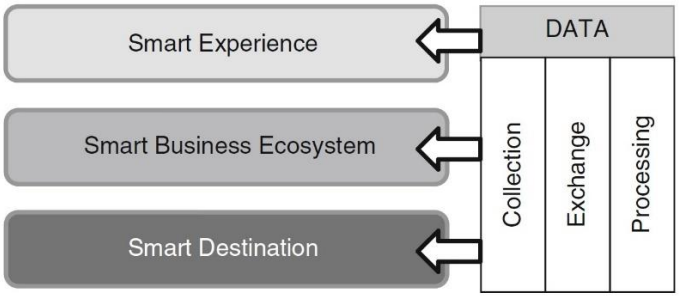
As the smart city concept encompasses a wide range of sectors, including tourism (Guo et al., 2014), some scholars, such as Jasrotia & Gangotia (2018), claim that smart tourism is an integral aspect of smart cities. The tourism industry has been experiencing rapid growth globally, apart from the downturn during the COVID-19 pandemic, and can be considered an important source of revenue for many cities, thus it should be paid attention to when planning a city's development (Mazanec & Wöber, 2010; Morrison & Coca-Stefaniak, 2020; UNWTO, n.d.-d). Therefore, the focus should not only be on developing infrastructure as well as improving the lives of permanent residents, but also on enhancing the travellers', and thus short-time residents, experiences (Jasrotia & Gangotia, 2018; Lamsfus et al., 2012). Urban tourism is at an all-time high and is expected to continue gaining importance (UNWTO, 2012). Nonetheless, tourism prevails to be under-managed in many cities

throughout the world, according to the WTTC's (2019b) study on cities' ability to keep up with the tourism industry's development. Overtourism has become a phrase that encapsulates the burden that tourism may have on urban infrastructures and shows inhabitants' dissatisfaction as a result of poorly managed tourism (Dodds & Butler, 2020; Gretzel & Koo, 2021; Milano et al., 2019). Smart tourism is inextricably tied to sustainability because of its relationship to smart cities, which arose from the necessity for sustainable development (Gretzel & Koo, 2021). However, it mostly refers to environmental sustainability rather than the triple bottom line of environmental, economic, and social sustainability (ibid.). Next to that, the tourism industry is confronted with several obstacles nowadays, including the emergence of disruptive forces such as technological advancements (Buhalis & Amaranggana, 2013; Buhalis & Law, 2008), which have fundamentally transformed the industry (Neuhofer et al., 2015). The rise of the idea of smart destinations is driven by the birth of the age of information (Jacobsen & Munar, 2012) and the emergence of interconnected tourists who are becoming increasingly digital as a result of the extensive use of mobile devices, social media and other forms of ICTs (Buhalis & Foerste, 2015; Gretzel, Sigala, et al., 2015; Jovicic, 2019; Xiang, Tussyadiah, et al., 2015). Consequently, changes in visitor behavior have occurred (Rotchanakitumnuai, 2017) as today's visitors are more connected, demanding and knowledgeable, particularly in terms of technological skills (Iunius et al., 2015), than ever before (Jovicic, 2019; Neuhofer et al., 2012). Therefore the drive of smart tourism comes not only from the tourism industry, particularly destination management organizations (DMOs), but also from tourists (Oates, 2016). Academics such as Jovicic (2019) and Jasrotia & Gangotia (2018) claim that it is nearly impossible to achieve successful market valorization or even to "survive" without the use of digital and smart technologies in the present day. Thus, both destinations and tourism businesses must be able to adapt their business strategy (Celdrán Bernabeu et al., 2016; Gretzel, Sigala, et al., 2015).

All smart tourism efforts revolve around the gathering, sharing, and processing of data (Xiang & Fesenmaier, 2017b). This may be obtained by destinations from physical infrastructure (e.g., through sensors), social networks, official sources, visitors and citizens, as well as modern technologies (e.g., Big Data and IoT) from inside and outside the tourism field to convert that data into business value propositions and improve strategic management which will ultimately aid in the enhancement of the visitor experiences and the provision of real-time and personalized services (Buhalis & Amaranggana, 2013; Choe & Fesenmaier, 2017; Gretzel, Sigala, et al., 2015; Sigalat-Signes et al., 2020). Tourism enterprises and destinations must be able to not only gather massive volumes of data, but also store, process, integrate, analyze, and utilize that data intelligently to create tourism activities, services, and innovations (Xiang & Fesenmaier, 2017b). The ultimate goal is to use this data to increase destination competitiveness and customer satisfaction, next to sustainability through enhanced resource management effectiveness all across the destination (Boes, Buhalis, et al., 2015; Buhalis & Amaranggana, 2013; Presenza et al., 2014; UNWTO, n.d.-a).

The smart tourism concept entails the three components of smart experiences, smart business ecosystems, and smart destinations (cf. Figure 2; Gretzel, Sigala, et al., 2015a). A smart information layer, which collects data; a smart exchange layer, which facilitates interconnectedness; and a smart processing layer, which is responsible for data analysis and visualization, support these three elements (Gretzel, Sigala, et al., 2015; Tu & Liu, 2014).

Figure 2: Components of smart tourism



Source: Gretzel, Sigala et al. (2015, p. 181).

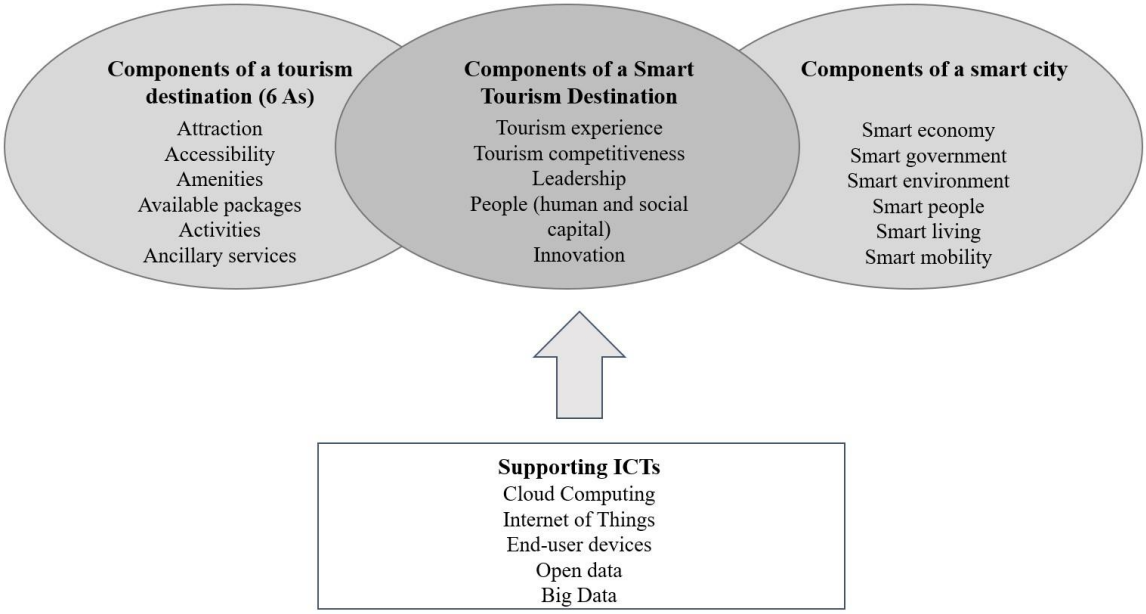
Smart experiences concentrate on technology-mediated travel experiences and how to improve them through value-added personalization based on tourist data, real-time monitoring, and feedback loops (Boes, Buhalis, et al., 2015; Buhalis & Amaranggana, 2015; Del Chiappa & Baggio, 2015; Neuhofer et al., 2015). Real-time responses are critical as they allow for fast problem resolution and prevent customers from leaving negative reviews (Buhalis & Amaranggana, 2015). Tourists are active actors in the experience co-creation as they not only consume but may also produce the data that forms the experience's foundation, e.g., by sharing pictures online (Gretzel, Sigala, et al., 2015; Neuhofer et al., 2012). Smart tourism also introduces a **smart business ecosystem** that includes novel resources, dynamically interconnected actors, and exchange methods (Buhalis & Amaranggana, 2013; Gretzel et al., 2016; Gretzel, Sigala, et al., 2015). Businesses from various industries and even outside the boundaries of a single city (e.g., international airlines) complement the smart economy and utilize local infrastructure and resources to deliver smart experiences (Gretzel, Werthner, et al., 2015). In smart tourism, tourists are recognized as temporary residents (Anuar & Gretzel, 2011; Gretzel, Sigala, et al., 2015). The third component of smart tourism is **smart destinations** which will be defined further in the subsequent section.

1.1.4 Conceptualisation of Smart Tourism Destinations

The term Smart Tourism Destination builds upon the foundations of a tourism destination, namely the six As, as well as the notion of smart cities (Avelar, 2020; Baggio & Cooper, 2010; Buhalis & Amaranggana, 2013; Cimbaljević et al., 2019; Del Chiappa & Baggio, 2015; Giffinger et al., 2007; Gretzel & Koo, 2021; Ivars-Baidal et al., 2019; Jasrotia &

Gangotia, 2018; Khomsi & Bedard, 2016; W. Zhu et al., 2014), but with a greater emphasis on tourism development and touristic infrastructure (Gretzel, Reino, et al., 2015; Sigalat-Signes et al., 2020). They make use of the ICTs, infrastructure, and innovations of smart cities where the tourism industry is one of the main sources of revenue (Avelar, 2020; Boes, Buhalis, et al., 2015; Jasrotia & Gangotia, 2018; Sigalat-Signes et al., 2020). While smart cities aim to improve residents' quality of life, smart destinations seek to also improve and co-create visitor experiences via the use of ICTs (Boes, Buhalis, et al., 2015; Caragliu et al., 2011; Neuhofer et al., 2012). Moreover, Buhalis and Amaranggana (2013) and Boes et al. (2016) claim that a smart destination's technologies are fundamentally different from those employed in a smart city since those are used *before, during and after* the visit to STDs while in smart cities the deployed technology is confined to usage *within* the city. This conceptualization can be summarized as shown in Figure 3.

Figure 3: Dimensions of a Smart Tourism Destination framework



Source: Own work (Adapted from Boes, Buhalis, et al., 2015; Boes et al., 2016; Cimbaljević et al., 2019).

A rising number of scientific studies have attempted to investigate the roots of this idea (Baggio et al., 2020) and came to the conclusion that there is no single definition of a STD yet (Tyan et al., 2020) as the notion itself is still evolving (Del Chiappa & Baggio, 2015). This also leads to challenges in finding a tourism destination that is completely smart (Baggio et al., 2020). While researchers define the term in a variety of ways, the majority of them emphasize the importance of ICTs and seem to imply, at least indirectly, that the ubiquitous usage of technology would allow a destination to be considered "smart" (Baggio et al., 2020; Tyan et al., 2020). Sensors, NFC, and radio-frequency identification (RFID) are just a few examples of the employed technologies (Gretzel, Reino, et al., 2015). These generate massive multidimensional sets of data, which are referred to as Big Data, and not

only enable DMOs to get useful insights (Marine-Roig & Anton Clavé, 2015) but also change the way tourists engage and relate with smart destinations (Ivars-Baidal et al., 2019). Rather than gathering and using data only from citizens, STDs also collect data from visitors (Buhalis & Amaranggana, 2013; Musa, 2016). The implementation of ICTs into physical infrastructure to collect information in real-time (López de Ávila Muñoz & García Sánchez, 2013) is a critical feature for the development and launch of smart destinations (Gretzel, Sigala, et al., 2015). This is important for attracting tourists (Lamsfus et al., 2012; Rotchanakitumnuai, 2017) and enhancing tourism experiences (Egger, 2013; López de Ávila Muñoz & García Sánchez, 2013; Rotchanakitumnuai, 2017) and destination competitiveness (Boes, Buhalis, et al., 2015; Buhalis & Amaranggana, 2013; Rotchanakitumnuai, 2017). It further improves the development, efficiency, and effectivity of tourism services and processes (Boes, Buhalis, et al., 2015; Egger, 2013; Wang et al., 2013).

Yet, while the use of ICTs is a vital facilitator for STDs, it can only assist a tourism destination's shift to smartness if the physical components of the ecosystem are substantially redesigned (Baggio et al., 2020; Boes, Buhalis, et al., 2015). STDs must be viewed as part of a larger, and thus more complex and dynamic, smart tourism ecosystem (Baggio et al., 2020; Del Chiappa & Baggio, 2015; Gretzel, Werthner, et al., 2015; Jovicic, 2019). Moreover, smart tourism should be regarded as a comprehensive tourism development and management philosophy with broader implications for its strategic orientation (Gretzel, 2018). Soft smartness components like leadership, strategic management, vision, innovation, and social capital are also essential components of smartness, in addition to hard smart components such as ICT (Boes, Buhalis, et al., 2015; Boes et al., 2016). Only the combination of hard and soft smartness components has the potential to provide the promised benefits and a long-term competitive advantage (ibid.).

A STD can thus be summarized as an innovative tourist destination founded on state-of-the-art technological infrastructure that ensures sustainable development and accessibility and allows tourists to co-create their environment, thereby improving the quality of the visitor's experience, the residents' lives as well as the destination's financial gains (Boes, Buhalis, et al., 2015; Lopez de Avila, 2015; López de Ávila Muñoz & García Sánchez, 2013; UNWTO, n.d.-a). The concept is the outcome of tourism destinations' connectedness with numerous stakeholder groups via technological platforms integrated with the environment and open innovation (Ardito et al., 2019; Buhalis & Amaranggana, 2015; Del Chiappa & Baggio, 2015; Jovicic, 2019). Using centralized data platforms, both the tourism sector and tourists benefit from having access to a wide array of information (W. Zhu et al., 2014). Smart destinations are an important driver of innovation, digitization, creativity, and competitiveness in tourism (Dexeus, 2021; Gretzel et al., 2016) and are essential to the tourist industry's transformation (UNWTO, n.d.-a).

Despite the numerous benefits provided by STDs, they also confront obstacles, like privacy concerns (Buhalis & Amaranggana, 2013) as STDs collect information about people, their actions, and position using intelligent technology, which might be regarded as a possible

danger (Vanolo, 2014). Thus, in order to examine not only their ability to benefit but also their potential to damage users, intelligent systems in tourism must be evaluated (The Economist, 2013). Another source of concern is access to data and real-time data exchange, which is critical in STDs. To achieve data openness in STDs, tourism authorities should assure that all information is made publicly available without additional expense, subject to their legal consent (Buhalis & Amaranggana, 2013). Furthermore, it is vital to emphasize that the technologically illiterate and the underprivileged citizens, tourists as well as business owners have serious limitations within destinations as those are held responsible to keep up with technological changes on their own (Buhalis & Amaranggana, 2013). Hence, it is advised that destinations focus not just on optimizing the use of new technology but also on educating tourism actors on how to make the most of it (Komninos et al., 2012). Lastly, Gretzel et al. (2016) argue that it remains unclear whether cities are truly prepared to take advantage of smart tourism's numerous benefits as there is no universal approach to becoming an STD. STD managers must comprehend the interconnectedness of the several basic components of smartness needed in order to become STDs (Boes et al., 2016).

When looking at the future development of the STD concept, some authors such as Gretzel & Koo (2021) claim that a distinction between smart cities and STDs is illogical since the boundaries between residential and touristic regions are progressively blurred through smart technology whilst synergies are created between them. For example, the establishment of free wi-fi meets the demands of both locals and visitors. Smart tourism cities would enable information sharing on a far larger scale than smart destinations by merging everyday and touristic applications of these technologies (Gretzel, Sigala, et al., 2015). This holistic concept focuses on the well-being of all people involved, rather than just targeting the quality of life of residents or the quality of experience for tourists. Other ideas include expanding smart tourism beyond individual smart destinations to smart tourism regions (Gretzel, 2018).

1.2 Big Data and data analytics

1.2.1 History and evolution of data analytics

Data analysis is the practice of examining raw data to make conclusions and uncover trends in order to improve processes and a system's overall performance (Xiang & Fesenmaier, 2017a; Y. Zhu & Xiong, 2015). At an ever-increasing rate, the modern globalized world has been exposed to continual and drastic social, political, and economic changes (Porter & Heppelmann, 2014). The adoption of cutting-edge technology is one strategy to handle these shifts (Morales-Urrutia et al., 2020; Townsend, 2013). The development of Information Technology may be broken down into three stages. First, the internet was invented, resulting in the creation of websites and global distribution channels (Xiang & Fesenmaier, 2017b). The experience economy, in which technology is utilized to improve experiences, emerged during the second stage (ibid.). Finally, search engines, social media, mobile technologies,

and the IoT made an appearance. Along with these latest advancements, data has grown to be ubiquitous. In 2021, Google executed more than 5.7 million searches every minute of the day, while TikTok users viewed 167 million videos and 575.000 tweets were published (Domo, n.d.). The volume of data used by modern applications is rapidly increasing, creating hurdles for distributed computing platforms to handle the massive amounts of data (European Commission, 2021; Kambatla et al., 2014; Katal et al., 2013). To put that into perspective, according to the World Economic Forum (Desjardins, 2019), the quantity of data in the world was estimated to be 44 zettabytes at the start of 2020, and by 2025, 175 zettabytes of data are expected to be in the international datasphere (Seagate UK, 2018). For reference, one zettabyte equals 1,000,000,000,000,000,000,000 bytes. These examples give an insight into the fast-developing scope and scale of varied sources of datasets that exist today (Kambatla et al., 2014).

Smart data usage may influence all sectors of the economy, opening up new avenues for economic growth, even for small- and medium-sized enterprises (SMEs) (European Commission, 2021). In 2019, the EU27 data economy was worth over €325 billion, accounting for 2.6 percent of its GDP, and is projected to exceed €550 billion by 2025 (European Commission et al., 2020). Within a few years, it is expected that data will be used in the majority of economic activity (McAfee & Brynjolfsson, 2012). Data is not only the core of the economy but also a driver of innovation since it serves as the foundation for the development of Artificial Intelligence (AI) and BDA (European Commission, 2021).

For a while, **Big Data** has been one of the most popular buzzwords (d'Amore et al., 2015; SAS, n.d.; Sheng et al., 2017), gaining particular traction in the first decade of the twenty-first century (Davenport, 2019). Businesses started employing fundamental analytics to identify insights and patterns as early as the 1950s, decades before the term "Big Data" was coined (SAS, n.d.). Xiang and Fesenmaier (2017b) describe the advent of BDA not merely as a buzzword but as a result of developments in computer engineering, adoption of IT, and the industry's quest for efficiency and novel methods of measuring productivity and performance, particularly in the last two decades. Most firms now recognize that if they collect all of the data that flows into their operations, they can apply analytics and derive tremendous value from it (SAS, n.d.). The word Big Data originally appeared in Bryson et al.'s (1999) foundational study. The concept refers to an enormous data set with a high degree of variety, velocity, and complexity, among other characteristics that will be covered in further depth in Chapter 1.2.2 (Davenport, 2019; Manyika et al., 2011). According to Google Trends, people are becoming increasingly interested in Big Data, especially after 2011/2. The values in Figure 4 indicate the search interest relative to the highest point on the chart, where the value 100 represents the highest popularity of the search term. Also scholars are fascinated by the concept of Big Data which is proven by the increasing results on Google Scholar. The academic search engine provides 5,280,000 hits for the phrase "Big Data" (as of February 2022), which cover different areas such as computer and social sciences.

Figure 4: Trends in worldwide Google searches for "Big Data"

Source: Google Trends (2022).

In the past few decades, increasingly complex platforms and systems for handling Big Data have been deployed, providing various benefits over old methodologies for aspects such as data management and analytics (Sheng et al., 2017). As Big Data grows in importance, new approaches will become more popular since these sophisticated tools have more functionalities, making corporate analytics faster, cheaper and more efficient (SAS, n.d.; Sheng et al., 2017). Following its rise in popularity, some researchers have regarded Big Data as a type of panacea, capable of solving societal problems and providing a plethora of beneficial insights into various areas of the lives of individuals, companies, and markets (Mayer-Schönberger & Cukier, 2013; McAfee & Brynjolfsson, 2012). Effective data utilization may assist in addressing a wide range of issues varying from medical purposes to environmental ones (European Commission, 2021). According to some academics, Big Data is the next "big thing" (Gobble, 2013), promising a revolution in science, technology, and productivity (Ann Keller et al., 2012; Manyika et al., 2011), as well as the ability to transforming the whole business process (Fosso Wamba et al., 2015).

1.2.2 Big Data definition and characteristics

Although the notion of Big Data has been known for a while, it is a constantly developing concept that changes with time and industries (Alastal & Badawy, 2020; Manyika et al., 2011). Given that the quantity of data continues to grow, there is no established criterion for determining what size and kind of data may be considered Big Data (Sheng et al., 2017) and no common definition of the term (Al Nuaimi et al., 2015; J. Li et al., 2018; Sheng et al., 2017). In its current usage, the phrase "Big Data" was established in 2005 to describe a massive volume of data that conventional data management approaches are unable to handle and analyse owing to its high complexity due to the diverse and complex formats of the data (Bibri & Krogstie, 2017; Gandomi & Haider, 2015; Halevi & Moed, 2012; Mariani et al., 2018; Marine-Roig & Anton Clavé, 2015; Oussous et al., 2018; Ularu et al., 2012).

Early in the development process of the Big Data concept, the three Vs of Big Data, namely **volume**, **variety**, and **velocity**, were introduced (M. Chen et al., 2014; Gandomi & Haider,

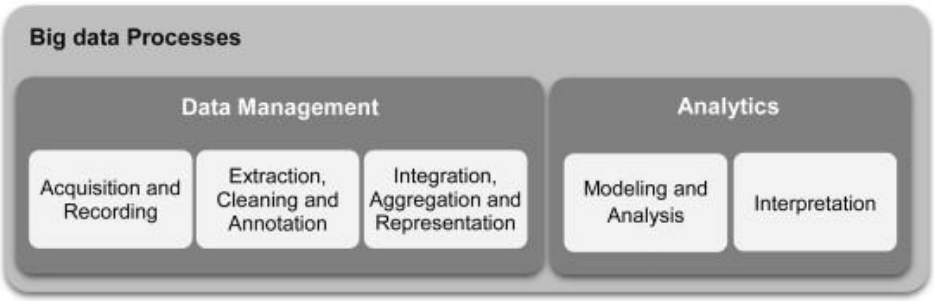
2015; Kwon et al., 2014; Laney, 2001; McAfee & Brynjolfsson, 2012; Russom, 2011; Song & Liu, 2017). These three attributes describe the progress of more and more data being generated and gathered at an accelerating rate from diverse sources such as texts (e.g., documents), multimedia data (e.g., pictures, audio, video), online data (e.g., websites), smart devices, mobile data (e.g., sensors, satellites and Geographical Positioning Systems (GPS) signals) and social media (Jin et al., 2015; Jones, 2012; Okwechime et al., 2018; Sheng et al., 2017). These are collected in near real-time, thereby increasing the complexity (Mariani et al., 2018; Russom, 2011; Sheng et al., 2017; Song & Liu, 2017). Data is not just structured, but also semi-structured and unstructured (Kambatla et al., 2014; Sheng et al., 2017). In fact, 80 percent of tourism-related data is unstructured, requiring the use of new technology to analyze it (Song & Liu, 2017). As for the velocity of tourism Big Data, there are three significant features (M. Chen et al., 2014). The first one is to acquire, store, and analyze the fast-moving flows of Big Data in a consistent and comprehensive manner (ibid.). The properties of timeliness or delay are the second factor to consider. Big Data should be acquired, kept, and used within a specified time frame, depending on the kind of information, as some data is permanent and other will become obsolete within a short amount of time (ibid.). The third consideration is the speed with which Big Data is stored and retrieved as real-time turnaround must be supported (M. Chen et al., 2014; Song & Liu, 2017). The fourth V, **veracity**, is defined by Baggio (2016), IBM (2014), and White (2012) as the accuracy, reliability, and validity of the data. Biases, abnormalities in data and noise caused by duplications are further aspects of veracity that should be avoided (Song & Liu, 2017). In comparison to volume and velocity, veracity is the most challenging task since the data must be maintained clean (Song & Liu, 2017). Some researchers went so far as to add two more characteristics to the definition of Big Data: **variability**, which describes the data's changing meaning across contexts and time, and **value** (Baggio, 2016; Pérez, 2015), which emphasizes Big Data's usefulness (J. Li et al., 2018) in terms of generating higher revenues, reducing costs or improving added value for customers. These latter three Vs originate from a management perspective and complement the original three Vs which stem from a more technological perspective (Gajdošík, 2019; Manyika et al., 2011; McAfee & Brynjolfsson, 2012). The seventh, and last, V is known as **visualization** (Baggio, 2016; El Alaoui et al., 2017) and addresses the requirement for trends and patterns to be presented in a graphically intelligible manner (Mariani et al., 2018).

To conclude, Big Data may be characterized as a comprehensive method for managing, processing, and analyzing the seven Vs (volume, variety, velocity, veracity, variability, value and visualization) in order to provide actionable insights for long-term value delivery, performance measurement, and competitive advantage (Fosso Wamba et al., 2015).

1.2.3 Process of Big Data analytics

According to Chen et al. (2014), there are four phases in the Big Data process, which can be summarized as data generation, acquisition, storage, and analysis. Gandomi & Haider (2015) divide the Big Data value chain into two phases instead, namely data management and analytics (cf. Figure 5) which will be elaborated further in the following subsection.

Figure 5: Big Data processes



Source: Gandomi & Haider (2015, p. 141).

Once the data has been gathered from various sources such as sensors, people, electronic data readers, and others, it must be stored, categorized, and processed in order to deliver relevant results (Al Nuaimi et al., 2015). One crucial aspect of the Big Data process is **data management**. Data needs to be of high quality and well-governed before it can be reliably analyzed (SAS, n.d.) through extracting, cleaning, and annotating the data (Gandomi & Haider, 2015). With data constantly flowing in and out of an organization, it is important to establish repeatable processes to build and maintain standards for data quality (SAS, n.d.). By filtering out the noise, the pace of making informed decisions can be accelerated (ibid.). Since data analytics typically require a lot of processing power, Big Data applications necessitate reliable and scalable software and hardware platforms (Al Nuaimi et al., 2015). Tools for BDA include Hadoop, MapReduce, HPPC, and Apache Storm, among others (Al Nuaimi et al., 2015; Rodríguez-Molano et al., 2018). Hadoop, for example, is a free open-source software framework that can store large amounts of data and run applications on clusters of hardware (SAS, n.d.). Big Data can also be stored and processed in the Cloud, relieving program owners of the obligation of acquiring specialized platforms, which are typically quite expensive, and instead allows them to utilise already tested, reliable platforms provided by Cloud service providers (Ji et al., 2012). Once the data has been collected, cleaned, and stored, it can finally be integrated with other data and represented to generate insights within the next step of BDA (Gandomi & Haider, 2015).

Big Data analytics is the act of examining large amounts of data using analysis algorithms to uncover hidden patterns, correlations and other insights almost immediately, therefore exploiting the value of data for decision-making (Alastal & Badawy, 2020; Gandomi & Haider, 2015; H. Hu et al., 2014; SAS, n.d.; Xiang & Fesenmaier, 2017b). Compared to traditional research and development methods, Big Data analytics improves human

capabilities to understand the consumer market at an unprecedented scale, scope, and depth (Boyd & Crawford, 2012). Several types of technology work together to help get the most value from data (SAS, n.d.). Most techniques can be classified as Machine Learning, a specific subset of AI that describes the methods and algorithms used for mining data and trains a machine on how to learn (Mitchell, 1997; SAS, n.d.; Witten et al., 2011). This makes it possible to quickly and automatically produce models that can analyze bigger, more complex data and deliver faster, more accurate results whilst extracting general rules, patterns, correlations and knowledge from unstructured data sources (ibid.). The last step in the process of Big Data is the interpretation of the generated results. Data can be used to describe what has happened, diagnose why it has happened, predict what will happen and even prescribe how one should react to a given situation by offering decision support or even automation (Tapadinhas, 2014). Predictive analytics may be used to assess the likelihood of future occurrences such as those related to fraud, risk, and marketing based on historical data which gives firms more confidence in their decision-making (SAS, n.d.).

1.3 Data in tourism

Over the past years, the tourism industry has been progressively growing and diversifying, making it a more globalised and competitive sector (Morales-Urrutia et al., 2020; Sigalat-Signes et al., 2020; Xiang & Fesenmaier, 2017a) whilst having continual innovation demands (Sigalat-Signes et al., 2020). It is thus the duty of tourist destinations' management to establish innovative and sustainable strategies (Sigalat-Signes et al., 2020). In the end, the implementation of ICTs heavily impacts innovations and transformations in the tourism sector (Buhalis, 2015; Choe & Fesenmaier, 2017; Hjalager, 2010; Koo et al., 2015; Morales-Urrutia et al., 2020; Sigalat-Signes et al., 2020; Xiang & Fesenmaier, 2017a) that are being adopted to enhance the administration, logistics, and marketing of touristic products and services (Morales-Urrutia et al., 2020). This emphasizes the need for enterprises to invest in the most relevant and latest technical solutions (Ruiz Garcia & Hernández García, 2017). In the following sections, the application of both ICTs and data analytics in general, and Big Data in particular, within the tourism sector will be elaborated.

1.3.1 ICTs and data analytics in tourism

Cutting-edge technologies such as Big Data, cloud computing, location-based services, virtual reality (VR), augmented reality (AR), and social networks are transforming consumer experiences and driving creative business models in the tourist industry (Gretzel et al., 2016; Wang et al., 2012). Modern technology may be utilized to empower not only the supply side of the tourism industry (i.e., destinations and travel enterprises) but also the demand side,

travelers, to personalize tourism products (Andrejevic & Burdon, 2014; Swan, 2013). Tourists are now well-informed thanks to the expanded use of ICTs and the subsequent interchange of information (Del Chiappa & Dall'Aglio, 2012). Since tourists nowadays regularly make travel choices using internet tools (Directorate-General for Communication, 2021), travelers are now able to adjust quickly the itinerary owing to the availability of information and the usage of mobile devices (Lamsfus et al., 2015; Xiang & Fesenmaier, 2017a). Smartphones, apps, and wearable gadgets such as the Apple Watch, have broadened the range of tourism experiences and allow visitors to connect and share their experiences (Gretzel et al., 2016; Tussyadiah, 2013; Wang et al., 2012, 2014). Travelers' decision-making behavior (Wang & Xiang, 2012), as well as the way they access and consume tourism products (Cimbaljević et al., 2019), have been found to be profoundly altered by today's technology. This requires the employment of new analytical tools that can accurately grasp travel behavior (Gretzel, Sigala, et al., 2015; Xiang, Schwartz, et al., 2015; Xiang & Fesenmaier, 2017a), which tourism management performance is becoming progressively reliant on (Xiang & Fesenmaier, 2017a).

ICTs can be applied in STDs in a number of ways, as Buhalis and Amaranggana (2013) highlighted in their research. For example, VR and AR might aid interpretation in terms of experiencing the digital environment at heritage attractions (ibid.). Next, vehicle tracking offers real-time data that could be utilized for planning reasons, such as traffic control, security, and lastly, sustainability by tracking and minimizing pollution and energy consumption (Buhalis & Amaranggana, 2013; Hernández-Muñoz et al., 2011). In this regard, Chen, Ardila-Gomez, and Frame (2017) investigated how intelligent transportation systems might help smart cities save energy. Barcelona, for example, offers public bicycles throughout the city and allows travelers to track their whereabouts using an app, thus promoting ecologically-friendly mobility (Bicing, n.d.). Another use of ICTs in STDs would be apps or tools providing an overview of all tourism offers, opening hours, automated language translation, and tourist assistance (Buhalis & Amaranggana, 2013). Again, Barcelona may be taken as an example where interactive bus shelters display not just bus arrival times but also tourist information (Gretzel, Sigala, et al., 2015).

The importance of digital solutions in the tourism sector is also recognized by the European Commission Directorate-General for Communications Networks, Content and Technology and the Industry Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (2022) which thus launched the Tourism Transition Pathway. This plan includes digital transition as five of its 17 priority action areas for making the European tourism ecosystem more digital and competitive. These five aspects are: (1) developing a common framework and single market for using and sharing data within the European tourism data space; (2) expanding digital information of tourism offers; (3) using virtual technologies for tourism services such as virtual museum visits; (4) real-time management to better manage tourism flows; and (5) enhancing digital capacities such as skills, funding, infrastructure, and mentorship. In addition to the Tourism Transition Pathway, other initiatives such as the

new Industrial Strategy (A New Industrial Strategy for Europe, 2020), the European data strategy (European Commission Directorate-General for Communications Networks, Content and Technology., n.d.-a), and the Digital Compass 2030 (European Commission Directorate-General for Communications Networks, Content and Technology., n.d.-b) also acknowledge the need for digital transformations within Europe. With 93% of tour operators now having a website and 59% offering online booking, the pandemic served as a digital catalyst (European Commission Directorate-General for Communications Networks, Content and Technology., n.d.-b). Despite this, tourism SMEs in the EU still show a rather low level of digitalization due to a lack of infrastructure, funding, skills, guidance, and policy assistance (Dredge et al., 2018).

When contemplating ICTs as a potential resolution to the difficulties of destination competitiveness, Baggio et al. (2020) emphasize the significance of adopting a more critical-driven mindset. Since the 1990s, warnings can be found in literature about the dangers of utilizing technological tools without first thoroughly revising the mechanisms or processes in question, and how this can result in low efficiency or economic growth despite large investments in ICTs (Brynjolfsson, 1993; Hammer, 1990). Furthermore, ICTs have failed to reach the promised benefits in some economies, particularly developing ones, due to a lack of expertise and skill in the use of these technologies (Morales-Urrutia et al., 2020). Following, it can be concluded that in order to be effective, ICTs must be tailored to each specific environment (Robinson, 2012).

1.3.2 Applications of Big Data in tourism

As mentioned previously, there is a tremendous amount of heterogeneous data derived from social networks, apps, sensors, and other sources that can be capitalized on in the Smart Tourism sector (Formia et al., 2015; Kontogianni & Alepis, 2020; Masseno & Santos, 2018). These technologies appear to be leading toward a more data-driven sensor society, in which each person leaves a large data footprint every day, which opens up new potential for commercial development (Andrejevic & Burdon, 2014; Choe & Fesenmaier, 2017; Swan, 2013). Travelers leave electronic footprints during all tourism-related activities, as well as their decision-making and communication processes prior to, during, and after travel, like trip planning, booking, and giving feedback on social media platforms or online questionnaires (Ardito et al., 2019; Fuchs & Höpken, 2011; Hendrik & Perdana, 2014; Y.-H. Hu et al., 2017; D. Li & Yang, 2017; Lu & Stepchenkova, 2015; Song & Liu, 2017). The abundant supply of free information (Gretzel, Sigala, et al., 2015) is either voluntarily offered, observed, or deduced from consumers' behavior (Masseno & Santos, 2018). Astronomy, finance, and healthcare are some of the cross-domain data sources which are significant for tourism (Alastal & Badawy, 2020) and provide helpful insights for a variety of stakeholders, including residents, visitors, governments, and businesses (Lim et al., 2018).

Using Big Data, the tourism sector can better capture, analyze, and understand visitor behavior, stakeholder demands, and the tourism market, which can then be utilized to adapt to changing conditions and tailor tourism offerings (Gajdošík, 2019; J. Li et al., 2018). The application of Big Data and deep learning algorithms can also allow finding dynamics based on vast dispersed yet linked sets of data (Alaei et al., 2019; Song & Liu, 2017). Combining sentiment value with other data, such as transportation or weather, may lead to the discovery of previously unrecognized patterns (Alaei et al., 2019). Another potential utilisation of Big Data in tourism is to aid destination managers in gaining a better understanding of the influence of seasonality on tourism and sustainability (Demunter et al., 2017; UNWTO, n.d.-a). Therefore, it can assist them in forecasting and managing tourist demand and flows based on real-time data in order to manage destinations better (ibid.). In light of the ongoing COVID-19 pandemic, some countries, like Spain, have established dashboards for air travel demand, where destination managers utilize Big Data to design tailored marketing strategies based on the recovery of important source markets (Gallego & Font, 2021). Hence, in volatile scenarios such as a pandemic or any other type of crisis, Big Data may be utilized to discover changes in demand during and after the event. Another example would be that a destination can infer insights on what types of accommodation are preferred by travellers, where to invest, in which markets to promote itself, or in which languages tourist information should be offered. Higher customer satisfaction, enhanced tourism experiences, and targeted marketing campaigns are all outcomes of the implementation of Big Data in a tourism context (Ardito et al., 2019; D. Li & Yang, 2017).

Big Data may also be utilized to create new, quicker, and more detailed statistics, which is critical in the hotel and tourism industry (Gandomi & Haider, 2015). Currently, most predictive tourism activities are grounded on official data supplied by national or local statistical agencies, or on surveys done with conventional methods, which are then published with months of delay (Mariani et al., 2018). The validity of official tourist data has been criticized in the literature due to the currency of data and the lack of standardization of data gathering and statistical estimating methods (Lam & McKercher, 2013). Another concern is that many tourists go unnoticed as additional types of travel become available (Baggio, 2016). Therefore, one viable approach for improving data quality is relying on the many traces people leave online when utilizing various technological platforms (Baggio, 2016). With the continuing breakthroughs in and lower costs of Big Data analytics, it promises to become even more popular in additional tourism destinations (Gretzel, Reino, et al., 2015). In any regard, Big Data has yet to be established in more DMOs and STDs, as it is now employed mostly in travel corporations like Walt Disney Co. or NH Hotels, among others (Davenport, 2019; INVATTUR & Territorio Creativo, 2015). Disney World, for example, provides guests with an RFID-enabled wristband that collects information on their location and spending throughout their visit (Walt Disney World Resort, n.d.).

1.4 Opportunities and challenges of Big Data

The following sub-chapter is seeking to highlight the benefits and difficulties that Big Data may provide to Smart Tourism Destinations as debated in current research. This will form the basis for a later discussion on the opportunities and challenges found within the analysis.

1.4.1 Opportunities

Big Data holds great promises (Pries & Dunnigan, 2015) which can be summarized into six categories, namely business operations, decision-making and value creation; destination management; personalization, prediction of demand and improved tourist experience; smart cities; and sustainability (cf. Figure 6).

Figure 6: Opportunities of Big Data

Source: Own work.

Business operations, decision-making and value creation

When it comes to exploiting Big Data in the business sphere, current research in academia and industry demonstrates that Big Data technologies provide considerable assistance in the discovery, creation and realisation of value (Sheng et al., 2017). As for value discovery, Big Data may assist in initiating organisational change. Current technologies enable the gathering, transmission, and analysis of enormous data sets, offering real-time and novel insights (De Filippi, 2015) into human behavior and corporate transactions (Fosso Wamba et al., 2015; McAfee & Brynjolfsson, 2012; SAS, n.d.; Verhoef et al., 2016). This makes it easier for businesses to react to changes in the environment and consumer demands (Pérez,

2015; SAS, n.d.; Sheng et al., 2017), resulting in new and original goods and services. Overall, Big Data allows for better-informed decision-making and companies being able to uncover more effective methods of marketing as well as operations (SAS, n.d.; Song & Liu, 2017). This could encourage businesses to pursue new business possibilities by implementing enhanced business models (Fosso Wamba et al., 2015; Mayer-Schönberger & Cukier, 2013; SAS, n.d.). Big Data may also be utilized to gather information in order to decrease risk in uncertain scenarios (Williams & Baláž, 2015). Especially in times of crisis and as a result of the COVID-19 pandemic, Big Data has become extremely significant in the tourism business (Gallego & Font, 2021). UNWTO (2020) states that having an empirical basis for decisions and strategies is critical and proposes investing in data and analysis to monitor and mitigate the impact on tourism development as well as accelerate recovery. Strategic decision-making, increased operational efficiency (Sheng et al., 2017) and marketing effectiveness (Ambrosini & Bowman, 2009; Buhalis & Foerste, 2015; Gretzel, Sigala, et al., 2015) are all part of creating more value from Big Data within a business. Lastly, this value has to be realised within business development. In terms of financial performance, Big Data provides overall higher profitability and cost savings when storing vast volumes of data (SAS, n.d.). Using Big Data analytics, retailers may enhance their return on investment by up to 20% (Perrey et al., 2013). Furthermore, the utilization of Big Data may increase customer satisfaction (SAS, n.d.) and deepen customer relationships (Kiron, 2013). Big Data is also more trustworthy since it is based on users' actual behaviors rather than claimed intentions or skewed responses to surveys (Song & Liu, 2017). In the end, Big Data gives companies the capacity to generate a competitive advantage (Buhalis & Amaranggana, 2013; LaValle et al., 2010; Manyika et al., 2011; Wang et al., 2013) and a higher organisational success (Sheng et al., 2017). This is enabled by lower research and processing time and faster time to market (Al Nuaimi et al., 2015; Fosso Wamba et al., 2015).

Destination management

Big Data is also being advocated as a new toolset for destination management (Xiang & Fesenmaier, 2017b). Tourism authorities and DMOs may take advantage of geo-referenced Big Data sources as they give a more comprehensive view of visitor flow patterns (D. Li & Yang, 2017), which are useful for capacity management (e.g., overtourism) and destination planning (Shoval et al., 2014). Tourists may now share their user-generated content (UGC), such as Tweets or images, with finer geo-referenced data thanks to the advent of smartphones with built-in GPS transmitters (Hawelka et al., 2014; Vu et al., 2015). Eurostat has also acknowledged the importance of data in monitoring tourism activities (European Commission, 2020b). Lastly, the data must then be analyzed in such a manner that it can be used to create value, i.e., through informed decision-making (Celdrán Bernabeu et al., 2016).

Personalization, prediction of demand and improved tourist experience

Within the Smart Tourism context, sophisticated data mining techniques and analytics can be used to create contextualized offerings in real-time based on tourist demands (Buhalis & Amaranggana, 2015; Buhalis & Foerste, 2015), or to co-create goods and services with tourists, and thus provide higher value and an improved smart tourism experience to them (Davenport, 2019; Gretzel, Sigala, et al., 2015; Neuhofer et al., 2012; H. Zhu et al., 2015). This is, however, contingent on destinations' capacity to gather data, deliver omnipresent connections, and synchronize data in real-time (Neuhofer et al., 2015). If this is the case, travelers have access to a wider range of information (Choe & Fesenmaier, 2017). Further, it enables destination and tourism managers to track the complete travel journey, from the minute a visitor thinks about the destination before they actually start planning their trip until the moment they return to their everyday lives and share their experiences (Choe & Fesenmaier, 2017). This data may be used to create detailed segmentations and adjust product design, marketing, and services to match those demands accurately (Fosso Wamba et al., 2015; Jani et al., 2014; Pries & Dunnigan, 2015). Thereby, customer-centric experiences can be created with the consumer at the forefront (Song & Liu, 2017).

According to research, Big Data sources may be leveraged to analyze tourist experiences (Lu & Stepchenkova, 2015), track travel routes (Ahas et al., 2008), and evaluate tourist spending patterns (Kozak & Sezgin, 2013). Other studies have illustrated the prospects of Big Data derived from tourist-generated content in designing marketing tools such as recommendation engines (Day, 2011; Ghose et al., 2012). This makes the discovery of unknown patterns and emerging trends possible, as well as gaining an insight into their behaviors, demands and preferences (Aiden & Michel, 2013; Sigala et al., 2012; Song & Liu, 2017). This knowledge can then be used to create relationships with travelers (Song & Liu, 2017) and helps to stay competitive (Ndou & Beqiri, 2015). The goal of leveraging tourism Big Data is to develop a real emotional connection between consumers and tourist industry partners in order to increase customer service significantly, as well as to be proactive (Song & Liu, 2017). Tourism is a fast-paced sector, which necessitates swift data analyses as every need must be met immediately in order to remain relevant to visitors (Song & Liu, 2017). Otherwise, marketing might not result in any sales (ibid.). Further research has shown the benefits of evaluating social media Big Data in terms of visitor habits (Xiang & Gretzel, 2010), given that tourism requires prediction and decision-making assistance (Gretzel, Sigala, et al., 2015; Miah et al., 2017; Song & Liu, 2017). By integrating web traffic data from destination marketing firms to estimate hotel demand, Yang et al. (2014) illustrate the value provided in terms of enhanced forecasting models, demonstrating a reduction in errors in comparison to more traditional methods. Predictive analytics with tourism Big Data utilized in feedback mechanisms provide a significant improvement over traditional human feedback (Song & Liu, 2017). It is thus frequently promoted as a panacea for businesses as in the tourism sector, feedback is critical in determining client preferences, providing great experiences, and attaining company growth (ibid.).

Smart city

Many governments have begun to use Big Data to aid the creation, management and sustainability of smart cities all over the world (Al Nuaimi et al., 2015). Big Data may lead to improved urban liveability and quality of life (Al Nuaimi et al., 2015; Jiménez, 2015; Rodríguez-Molano et al., 2018) as a consequence of better-designed living and working spaces, better and quicker services, and, once again, the availability of sufficient information to make educated decisions (Al Nuaimi et al., 2015). The use of Big Data analytics is critical in a variety of fields (Gajdošík, 2019). They serve as the foundation for intelligent transportation systems (Al Nuaimi et al., 2015; Alastal & Badawy, 2020; Z. Khan et al., 2013; Xia et al., 2016) through public transportation routing, scheduling and traffic control (Lim et al., 2018), smart healthcare (Al Nuaimi et al., 2015; Alastal & Badawy, 2020; I. Hashem et al., 2016; Z. Khan et al., 2013; Pramanik et al., 2017), urban planning (Alastal & Badawy, 2020; Babar & Arif, 2017) and education (Al Nuaimi et al., 2015; Alastal & Badawy, 2020). Big Data is also used for governance and economic development (Alastal & Badawy, 2020; I. Hashem et al., 2016; Z. Khan et al., 2013), safety (Alastal & Badawy, 2020; Z. Khan et al., 2013), and public participation and information (Z. Khan et al., 2013), next to sustainability (Lim et al., 2018).

Sustainability

In terms of sustainability, Big Data may be used to track and allocate resources more efficiently, as many resources are becoming scarce or costly (Al Nuaimi et al., 2015). Hence, it is critical to integrate systems to identify waste areas and better distribute resources while managing costs and lowering energy and resource consumption in order to achieve more regulated resource usage (ibid.). Further, Big Data may help detecting possible environmental risks such as blizzards or monitoring air pollution (Lim et al., 2018) and other meteorological information which is of interest to agriculture (Al Nuaimi et al., 2015). In the tourism context, however, whilst economic performance is frequently measured, such as in tourism satellite accounts (Crouch & Ritchie, 1999), triple-bottom-line (namely, economy, environment and society) accounting at the destination level is mostly absent and has not been properly defined for smart tourism development (Gretzel & Koo, 2021). Yet again, given that Big Data originates from a number of sources, BDA might be a viable solution to defeat this (ibid.).

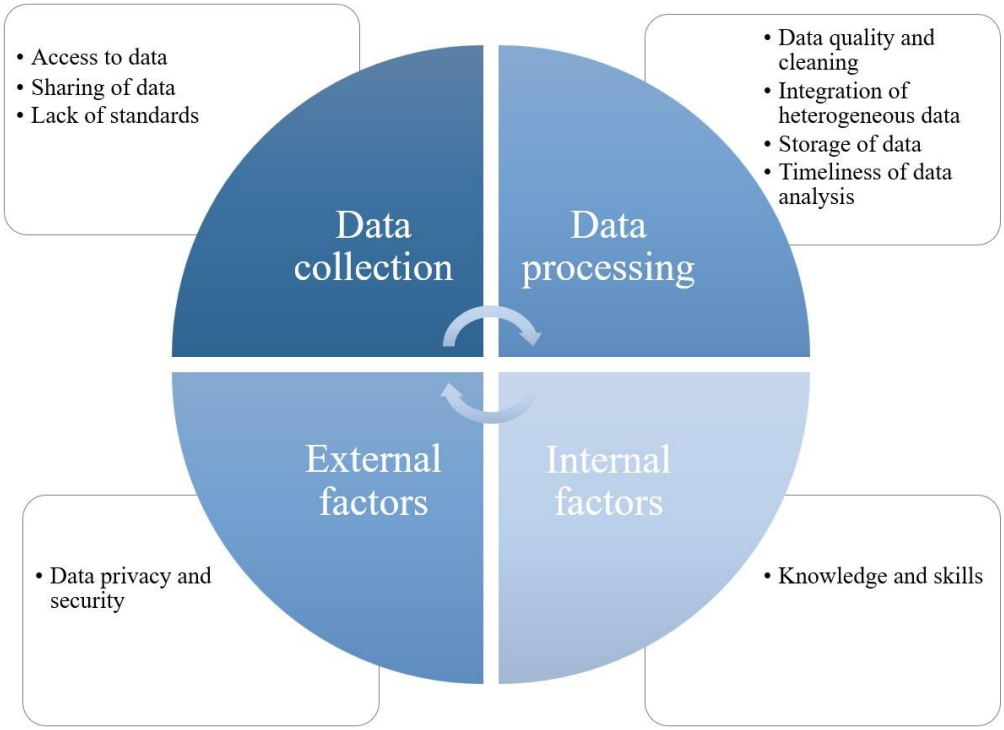
1.4.2 Challenges

The preceding sections demonstrate that Big Data offers several benefits to businesses, the tourism industry, and STDs in particular. However, like with any large-scale change

endeavor, the data-based transformation is not always straightforward (Lim et al., 2018), and difficulties must be addressed early on (Katal et al., 2013). Therefore, the issues that emerge during the transition of STDs utilizing Big Data will be discussed in this section.

Several studies (Boyd & Crawford, 2012; Ekbia et al., 2015; J. Fan et al., 2014; George et al., 2016; McFarland & McFarland, 2015) have addressed a variety of issues and threats that arise with Big Data. It presents challenges in terms of extraction, search, analysis, storage, transmission, visualization, and information privacy (Erl et al., 2016; I. A. T. Hashem et al., 2015), which have been summarized in Figure 7 and will be elaborated upon hereafter.

Figure 7: Challenges of Big Data



Source: Own work (Adapted from Lim et al., 2018 and Alastal & Badawy, 2020).

Data access and sharing

Getting access to data and exchanging information across sources represents the first challenge (Al Nuaimi et al., 2015; Gorini, n.d.; Komninos et al., 2012; Manyika et al., 2011). Though plenty of data can be found in most cities it is generally dispersed among several organizations since each government department often has its own storehouse of sensitive data that is not being shared (Al Nuaimi et al., 2015; Gorini, n.d.). Furthermore, requiring corporations to provide data about their clients and business is problematic since it jeopardizes the culture of discretion and competitiveness (Katal et al., 2013). When data is being utilized to make fast, accurate decisions, it must be provided in a comprehensive, precise and timely fashion (Katal et al., 2013). Thus, data must be made open and barriers ought to be removed in order to accomplish frictionless information sharing among various

institutions (Katal et al., 2013; Komninos et al., 2012; Su et al., 2011). Additionally, currently Big Data cannot be exploited to its greatest capacity because many companies lack a consistent method of data collection (Gorini, n.d.). Hence, data should be made accessible to all stakeholders in a uniform way, using standardized formats, resulting in improved decision-making, business intelligence, and productivity (Katal et al., 2013; Lee et al., 2013). The UNWTO is currently developing a uniform statistical methodology to address this problem (UNWTO, n.d.-c). The challenges in gaining access to data can also be expanded to the general public, where the digital gap is widening (Minghetti & Buhalis, 2010) for individuals without smartphones and places that cannot afford to establish smart tourism infostructures (Gretzel, Sigala, et al., 2015). Likewise, some inhabitants and visitors are hesitant to accept technology because they wish not to use it (Gretzel, Sigala, et al., 2015).

Data privacy and security

Privacy is becoming increasingly important in the Big Data world and should be given more attention (McAfee & Brynjolfsson, 2012). In particular, maintaining anonymity and privacy of personal information, as well as security and intellectual property, is critical (Al Nuaimi et al., 2015; I. Hashem et al., 2016; M. A. Khan et al., 2014; Kontogianni & Alepis, 2020; J. Li et al., 2018; Manyika et al., 2011). According to the WTTC's (n.d.) paper on Big Data and tourism, Big Data must focus on the consumer's experience without being too intrusive. There is a thin line between customers finding increased data collection invasive, resulting in a loss of loyalty, and consumers accepting those intrusions if the final product is more tailored and their needs are being anticipated throughout (M. A. Khan et al., 2014; Song & Liu, 2017). Consumer privacy remains critical (Gorini, n.d.), and confidential data such as customers' names, addresses, and financial information should be kept private and not shared with other parties (Bose, 2009). Certain data may also be subject to particular privacy restrictions, making it difficult to transfer between companies (Al Nuaimi et al., 2015; Mariani et al., 2018). In case databases store sensitive information, high-level security regulations and processes are required to secure this information against unauthorized access (Al Nuaimi et al., 2015). Furthermore, because data may travel over a variety of networks, some of which may be insecure, smart applications linked across institutions also necessitate a high level of security (M. A. Khan et al., 2014). What makes this situation even more complicated is that most large data platforms today, like Cassandra and Hadoop, are insecure (Kim et al., 2014). Further, while location-based services are tremendously beneficial to visitors, they often put customers' privacy at risk (Anuar & Gretzel, 2011). Indeed, the European Union and other regulating agencies have pressured numerous data-related companies, including Google and most telecommunications companies, to preserve customers' privacy rights (Xiang & Fesenmaier, 2017a). Otherwise, privacy issues hold tourism stakeholders (such as visitors, online travel agents, and hotels) back from disclosing personal information (J. Li et al., 2018). The new EU General Data Protection Regulation

(GDPR) emphasizes the importance of this need since data handling must adhere to a number of stringent guidelines (Proton Technologies AG, n.d.).

Data processing, integration and storage

The transformation, integration and processing of a multitude of heterogeneous variables, such as biological, physical, and social ones, into useful information presents another challenge (Al Nuaimi et al., 2015; I. Hashem et al., 2016; Komninos et al., 2012; Lim et al., 2018; Sun et al., 2016). Data processing has become more challenging as a result of constant technical advancements and the large growth in connected devices in metropolitan areas, resulting in a growing complexity of data sources and a greater range of data formats (Al Nuaimi et al., 2015; Kambatla et al., 2014; Mariani et al., 2018; Rodríguez-Molano et al., 2018). Many of the new data formats are unstructured, such as photos or audio, and must be managed and categorized into a structured format (Michalik et al., 2014), as the unstructured nature of the data makes it difficult to classify and organize in a way that organizations can make use of (Al Nuaimi et al., 2015; Mariani et al., 2018). When data comes from both inside and outside the company, problems with data integration and quality emerge (Mariani et al., 2018). Another difficult task is extracting valuable information from data by filtering the most significant data from all the data obtained by the company (Gretzel, Sigala, et al., 2015; Katal et al., 2013; Mariani et al., 2018; Sheng et al., 2017). Furthermore, if large amounts of data are not under control, they will become increasingly ambiguous and difficult to comprehend (Rodríguez-Molano et al., 2018). Therefore, hiring technically and analytically skilled data scientists is required to clean and organize the massive amounts of data (McAfee & Brynjolfsson, 2012). Aside from that, data storage is an additional difficult issue that comes with Big Data (Katal et al., 2013; Manyika et al., 2011; Mariani et al., 2018). The quantity of storage provided is mostly insufficient to store the amount of data created (Kambatla et al., 2014; Katal et al., 2013). Uploading it to the cloud will not address the problem as these terabytes of data will take a long time to transfer to the cloud, and since data changes quickly it will be difficult to upload it in real-time (Katal et al., 2013).

Data quality

Managing the quality of data is another essential element of Big Data in order to offer value to businesses, but it may also cause problems if not properly managed (Al Nuaimi et al., 2015; Fosso Wamba et al., 2015; Lim et al., 2018; Miele & Shockley, 2013). Otherwise, it may lead to misleading and faulty correlation when combining data which may result in the company making an inaccurate interpretation of a business opportunity (White, 2012). Poor quality or incorrect data has no prospects of assisting managers in making the best decisions; instead, it will waste organizational resources (Fosso Wamba et al., 2015). Multiple identical databases may provide inferior quality data, increase data storage costs and make data more

difficult to access and utilize (Beath et al., 2012). Duplicate, erroneous, and redundant data are always a threat, resulting from a lack of structure and consistency (Al Nuaimi et al., 2015; Fosso Wamba et al., 2015). Similarly, the credibility of online textual data has to be questioned, as some visitors may leave fake or fraudulent evaluations (Miele & Shockley, 2013). It is worth noting that the usage of even the most powerful analytical system is useless if the data is incorrect or of low quality (Bose, 2009).

Knowledge and skills

Many cities, governments and stakeholders in general are unsure of what to do with Big Data (Gorini, n.d.). Big Data in tourism is "a work in progress," according to a Skift research from 2017, which noted that data - even when available - does not always cascade down to local decision-makers (ibid.). Part of the reason might be that the data is frequently too complicated to comprehend. Organizational leaders frequently lack a grasp of the usefulness of Big Data, how to access it and extract insights from it (Manyika et al., 2011). As a result, the data insights cannot be translated into action plans (Gorini, n.d.). As Big Data is a new technology, it must attract businesses and workers with a wide range of novel skills (Katal et al., 2013). Business analytics abilities are still limited to the expert level, according to Shah et al. (2012), and have not yet been spread to all levels of an organization. However, in order to derive value from Big Data, all employees must be well-versed in the subject, which may be accomplished through training (Katal et al., 2013; Shah et al., 2012). Furthermore, higher education must implement a Big Data curriculum in order to develop qualified professionals (Katal et al., 2013). These abilities should not be restricted to technical abilities but should also include analytical, research, interpretative, and creative abilities (Katal et al., 2013). As a response, managers should arrange Big Data in a comprehensible manner, such as through reports, dashboards or other visualisation systems, in order to obtain more acceptability from workers (Bose, 2009; Gorini, n.d.).

Tourism

Many of the difficulties with Big Data in general apply to the tourism industry as well. Tourism destinations are experiencing a number of issues as a result of the increasing use of technology in infrastructure and the adoption of smartness (Soteriades, 2012). As previously stated, a hurdle in Smart Tourism is that data volume is always growing, necessitating suitable storage (Kontogianni & Alepis, 2020) as well as data filtering to remove the noise from abundant data (Braunhofer & Ricci, 2017; Nguyen et al., 2017). Furthermore, despite the fact that it is advised that DMOs construct mixed-source national tourism statistics that include surveys and Big Data, there is evidence that governments are unwilling to improve their existing tourism statistical methodologies with Big Data models (Demunter et al., 2017). Tourist surveys are presently utilized as the primary source of data for developing

national and regional tourism statistics (Saluveer et al., 2020) which serve as the basis for decision-making on a governmental basis (Lozano-Oyola et al., 2012). However, those are expensive and gather data retroactively despite being in times of fast change (Gallego & Font, 2021). Currently, the market for Big Data sources is very fragmented, with several private data providers, such as the telecommunications or banking sectors, that can give accurate tourism data (Demunter et al., 2017). This is a handicap for tourism destinations, which face technological and financial challenges in getting the information they need (Gallego & Font, 2021). Despite the fact that Big Data may be utilized to enhance traditional statistics and improve accuracy (Batista e Silva et al., 2018; Demunter et al., 2017), it is still considered unusual. To prevent duplication of effort, tourism authorities at the national, regional, and municipal levels should develop a standard framework for cooperating with data suppliers (Gallego & Font, 2021). They should also insist that government statistical agencies become involved in the utilization and distribution of these new sources of data (Gallego & Font, 2021). This strategy will not only increase data availability but also lower the cost of tourism research utilizing Big Data (J. Li et al., 2018). In this regard, the European Commission's (2020b) initiative, which has secured an agreement on data exchange through Eurostat with Airbnb, Booking, Expedia Group, and TripAdvisor, should be mentioned. The adoption of methodological norms, which have yet to be established, would facilitate cross-destination comparisons and ensure the process' transparency as well as the reliability of its outcomes (Gallego & Font, 2021).

Another issue is the shortage of data analysts in DMOs, as well as the requirement to train workers in new capabilities (Struijs et al., 2014). Despite its high reliance on ICT, tourism does not yet attract a large number of knowledge workers (Gretzel, Sigala, et al., 2015), and it also struggles with innovation (Hjalager, 2002). Top management teams in big hospitality organizations, such as hotel chains or airline corporations, require skilled data analysts who are up to date on the newest data science advancements (Davenport & Patil, 2012). Small and medium-sized tourist businesses that cannot afford to hire data scientists might purchase reports produced by Big Data analytics consulting firms (Mariani et al., 2018). Previously, DMOs concentrated all of their efforts on direct Big Data acquisitions, but they spent little time creating an action plan for what to do with the data after it arrived (Gallego & Font, 2021). It is vital to remember that having data is just not enough; the value of data comes from its capacity to effectively assist management and decision-making (Gallego & Font, 2021). To accomplish so, both human and artificial intelligence are required (Gretzel, Sigala, et al., 2015). While technical DMO employees may be able to develop useful insights using Big Data, there is a lack of common ground between them and policymakers who are intended to make choices based on such data, which poses hurdles to Big Data implementation (Nutley & Davies, 2000). Furthermore, Big Data academic programs for tourist management have been virtually absent (Mariani et al., 2018). As a result, academic institutions should invest in them (Mariani et al., 2018). Finally, it is important to note that merely integrating Big Data into a tourism destination is not enough to make it a Smart Tourism Destination (R. G. Hollands, 2008; Rotchanakitumnuai, 2017; Shapiro, 2006).

1.5 Future outlook of Big Data

Given its wide scope, all of the previously mentioned aspects are implicitly or explicitly related to the tourism sector. Big Data has enormous potential in tourism, and tourism businesses should not overlook its significance (Pries & Dunnigan, 2015). For tourist businesses and the industry as a whole, BDA may provide needed business insights and be a source of innovation (Song & Liu, 2017). According to Wang et al. (2013), these new insights may influence markets and organizations, as well as enhance corporate decision-making and the optimal allocation of resources. Big Data collected in visitors' daily lives and during travel, as well as possible business intelligence developed based on this data, are expected to serve as the foundation for the creation of STDs (Choe & Fesenmaier, 2017). With this in mind, it is suggested that the tourism sector is on the cutting edge of a new revolution that would alter not just the tools used to organize travel and the way travel experiences are constructed but also the tourism sector's very essence (ibid.).

Despite the fact that business managers and scholars are increasingly interested in Big Data, the full advantages of improving technologies and changing business settings have yet to be realized (Sheng et al., 2017). Firms must develop and implement a data-driven corporate culture, strategy, and skills to realize the maximum benefit while remaining adaptive to the ever-changing environment (McAfee & Brynjolfsson, 2012; Sheng et al., 2017). Therefore, investments in IT and the development of data analytics expertise are needed (Bertot & Choi, 2013; Frederiksen, 2012; Sheng et al., 2017). Finally, high-quality data and privacy must be guaranteed (Beath et al., 2012; Bertot & Choi, 2013). Executives will then be able to make smart, fast, and flexible decisions with the help of new approaches for analyzing and interpreting massive data (Sheng et al., 2017). Big Data can play a significant role in the development of nations, destinations and academic research (Jin et al., 2015).

As data gathering, information generation, and information dissemination are all interdependent tasks, it is vital to note that all of the aforementioned difficulties are linked (Lim et al., 2018). The challenges show that using Big Data for smart destinations demands a variety of skills, data management, and analytics (ibid.). Despite the obstacles mentioned above, scientific and corporate societies agree that Big Data has the potential to make a difference since it collects real-time behaviors and perceptions on nearly every element possible (Chang et al., 2014; Power, 2014). According to Fosso Wamba et al. (2015), the Big Data revolution is advancing, and businesses should embrace it in order to develop superior capabilities that may become a competitive advantage. Furthermore, the cost of Big Data analytics platforms is decreasing, and employees are getting more acquainted with what Big Data can accomplish (Song & Liu, 2017). Big Data has much promise for businesses in the future, especially in the tourist industry, if appropriate data sources are picked, workforces are trained, data is shared, cleansed, and useful knowledge can be retrieved.

2 METHODOLOGY

There are several approaches to investigating a certain research topic, depending on the researcher's viewpoints and convictions (Tracy, 2013). The entire research is guided by a selected paradigm which is why it is critical to discuss how the researcher understands and perceives the world and what is their methodological stance (Denzin & Lincoln, 2005). The latter is influencing the methods utilized to collect and analyze the empirical data (Tribe, 2001; Usher, 1996). Therefore, the selected research paradigm, the methods for collecting the data, as well as the methodology for the analysis will be outlined in the following section.

2.1 Research philosophy

A paradigm is a collection of beliefs that influence research, a specific perspective on how the world is experienced, understood and functions, and a professional devotion to the same regulations and principles for scientific conduct (Guba & Lincoln, 1994). Within research, one may apply several paradigms such as positivism, post-positivism, critical theory or constructivism, among others (ibid.). However, it is not a question of picking the appropriate paradigm but rather of following the paradigm that is most suited to the subject at hand (Guba, 1990; Kuhn, 1970). To address the aforementioned research question, the author will employ a interpretative/social constructivist paradigm, which states that there is no objective reality, implying that all perspectives and beliefs are socially constructed (Guba, 1990). According to social constructionism, people's perceptions of reality are always changing as a result of their experiences, therefore impossible to fully comprehend (Kvale & Brinkmann, 2015). This paradigm explains the researcher's approach to analysis through their understanding of the topics of Big Data and STDs (Rehman & Alharthi, 2016). As for this study, this is particularly significant as it delves into the socially created concept of Smart Tourism Destinations and man-made technological advancements.

The nature of reality is the subject of ontology (Denzin & Lincoln, 2005; Easterby-Smith et al., 2015; Guba, 1990; Jennings, 2015). One can differentiate between realism, relativism, and nominalism within the social sciences (Easterby-Smith et al., 2015). However, in the social constructivism paradigm, the ontology is relative, suggesting that there is no such thing as a fixed reality and its understanding and perceptions are dependent on the researcher's background, past knowledge and previous experiences (Hautamäki, 2020). As a result, the presence of multiple realities generated by various actors is recognized (Jennings, 2015), such as the diverse stakeholders to be investigated in this study. For this research, informants with different background knowledge have been contacted, meaning that they might have different approaches towards the phenomena being studied. Consequently, there is not necessarily one objective truth to the research question.

Epistemology deals with knowledge and defines the relationships between researchers and the subject (Denzin & Lincoln, 2005; Easterby-Smith et al., 2015; Guba, 1990). Within the social constructivism paradigm the epistemology is inter-subjective since the researcher's personal backgrounds and beliefs are used to construct knowledge (Guba, 1990). Given that the researchers had previously visited Valencia as a tourist, their epistemology is value-laden (Jennings, 2015), and their educational and cultural background might influence how they interpret the query. In other words, the researcher's understanding, as well as the understanding of the selected interviewees, contribute to the understanding of reality.

Lastly, the utilization of a constructivist paradigm enables the usage of a qualitative methodology as it seeks to gain a clearer understanding of the meaning dimensions of human interaction and social phenomena (Johannessen et al., 2020; Tanggaard & Brinkmann, 2015), which will be elaborated upon further in the following section.

2.2 Data collection

Research can be conducted using qualitative, quantitative or mixed methods (J. Creswell, 2003; Jennings, 2015). Qualitative methods are "used to understand people's beliefs, experiences, attitudes, behavior, and interactions" (Pathak et al., 2013, p. 192). They come with the advantage of giving a voice to the participants by letting them share their experiences (G. Gibson et al., 2004; Pathak et al., 2013), thus delivering more detailed results and understanding. On the other hand, it also comes with downsides such as potential subjectivity and biases of the researcher (Mruck & Breuer, 2003; Wells & Lo Sciuto, 1966; Wise, 2018). Moreover, qualitative methods are considered time-consuming and demand more resources for effective results (Ejimabo, 2015; Wells & Lo Sciuto, 1966) as they are less structured than quantitative ones (Mruck & Breuer, 2003).

This research is exploratory and descriptive as the theoretical concepts of Big Data and Smart Tourism Destinations are integrated and applied through the case study of Valencia, Spain. Exploratory research is usually employed when there is only little knowledge and data about the topic (Jennings, 2015) and is conducted to find preliminary indicators on the overall nature of an issue, to gain information in order to discover innovative approaches, and to determine decision-making alternatives (Malhotra, 2008). Due to the exploratory and timely nature of this research, a qualitative case study methodology was endorsed (Yin, 2009), which is widely used in tourism (Beeton, 2005) and is frequently utilized when investigations are still in their initial stages (Benbasat et al., 1987).

First, evidence was gathered regarding the development state of the smart destination using secondary research by reviewing sources of scientific and touristic information, as well as statistical sources from agencies and official public entities. Next to that, seven semi-structured in-depth interviews with key stakeholders in the city's tourism industry, and particularly in the field of smart tourism, technology and Big Data analysis, were held

between March-April 2022. These were used to develop conceptual categories that would detect and aid in comprehending the attitudes and perceptions of respondents toward the utilization of Big Data in the scope of Smart Tourism Destinations.

When deciding which research method to employ, one has to choose the one that is the most appropriate to attain the research aims and that aligns with the researcher's paradigms (Bagdonien & Zemblyté, 2005). Structured, semi-structured, and unstructured interviews are different types of interviews that are chosen based on the study issue (Magaldi & Berler, 2020; Tanggaard & Brinkmann, 2015). A semi-structured interview (SSI) was utilized in this study as it is feasible to maintain a more open discourse with the interviewees while also being able to ask follow-up questions and explain ambiguities (Adams, 2015; Johannessen et al., 2020; Kvale & Brinkmann, 2015; Magaldi & Berler, 2020). Moreover, it is the most common method of data collection in qualitative research (Sandelowski, 2000). According to Adams (2015), SSIs are a useful method to investigate unpredictably occurring issues, but also come with certain limitations, such as the need for skilled interviewers and it being a lengthy process. Since the respondents may have biased opinions on certain subjects, the legitimacy of the replies must also be considered. Next to that, an interview guide was created to provide a general framework and ensure that the most relevant subjects were addressed while allowing for additional spontaneous questions to be asked (cf. appendix). This also contributed to minimizing the interview method's potential flaws, such as concealed biases and the urge to "properly" answer questions (Phellas et al., 2011).

Given that STD is a relatively new concept with only little literature (Albino et al., 2015; Boes, Buhalis, et al., 2015; Meijer & Rodríguez Bolívar, 2015), the expert interview was found to be the most suitable approach for the purpose of this thesis as it is commonly applied within qualitative research to gather data (Mehmetoglu, 2004). The in-depth interview method is a standard approach for producing knowledge through the interplay between an interviewer and an interviewee in order to understand more about a phenomenon that is hard or even impossible to observe (Easterby-Smith et al., 2015; Kvale & Brinkmann, 2015). This method is particularly well-known for being a dialogue guided by the researcher (Yin, 2009), with participants exchanging viewpoints (Kvale & Brinkmann, 2015) and therefore was chosen as the data gathering approach because it allows participants to express themselves more freely (Johannessen et al., 2020; Phellas et al., 2011), without being impacted by other participants or social pressure, as it may be the case with focus groups (Easterby-Smith et al., 2015). Further advantages include the opportunity for the interviewer to respond to nonverbal signs, such as body language, and the duration of the interview being less regulated depending on the length of the answers (Phellas et al., 2011). Nonetheless, one of the interviews was conducted jointly, meaning that two respondents were interviewed at the same time, which may lead to one interviewee overpowering the other, and there is an underlying risk of inciting disagreements between the respondents (Arksey, 1996). The creation of more thorough data, provoking similar and different understandings, and giving a voice to individuals who might otherwise be silenced, on the other hand, are all potential

benefits of joint interviews. Overall, one could say that it is important to be aware of the disadvantages of the various research methods, but it does not mean that they are necessarily negative. In the end, most research methods have their limitations, some of which the interview method eliminates (Wells & Lo Sciuto, 1966).

The in-depth interviews were conducted in English, lasted between 32 to 97 minutes with an average length of 62 minutes and were held online over the video conferencing platform Zoom, given the current pandemic situation as well as the physical distance between the interviewees and the researcher. However, one of the interviews was conducted in written form over email due to time constraints and language barriers. Email interviews can be done asynchronously with participants from all around the world without incurring travel costs (L. Gibson, 2010; Hawkins, 2018). Participants can also react to email interviews at their leisure, giving them time to think about their replies before composing written responses. Yet again, writing those responses may take longer than answering orally, which leads to further downside of email interviews such as the potential for short responses, the possibility that participants drop out of the conversation and misinterpretations resulting from the limited interaction between researcher and the interviewee. Follow-up questions were sent to clarify ambiguous answers and attain more detailed responses; however, they were not responded to. The interview questions were provided to each of the experts before the interview, allowing them to prepare for the question as there was no need for a spontaneous reaction during the interview.

For analysis purposes, the interviews were recorded after the informed consent of the participants and were finally transcribed to ensure that no valuable information got lost and to facilitate data analysis. Transcriptions serve as the foundation for data analysis, allowing researchers to gain a deeper understanding of the information retained from the interviews as well as the identification of novel ideas (Ezzy, 2013). The majority of the spoken language was preserved, although certain interjections were omitted, and some phrases were rewritten in a more readable fashion. Moreover, certain referrals that might have been non-understandable to the readers were elaborated in square brackets, whilst abbreviations were explained in round brackets. As neither the researcher nor the research participants are native speakers of the English language, one must be aware of possible misunderstandings due to problems in expressing one's thoughts in a second language and because of accents. For this study, ethical concerns included informing all participants about the topic and goal of the interview, following which they all volunteered to participate in the study. Furthermore, all respondents were given the option of anonymity if they so desired.

2.3 Interview design

A well-designed interview is of utmost importance for academic research as a flawed interview design will lead to flawed collected data. The interview guideline was consisted

of 17 open-ended questions and can be found in appendix 2. These were created to ensure consistency in carrying out the study objectives. Solely open-ended questions were asked as they come with the advantage that respondents can openly express their views (Adams, 2015). Throughout the interviewing process, minor adjustments to the interview guide were made as questions were modified or put on a different focus throughout the interview, given that the researcher acquired more knowledge regarding the topic (Patton & Cochran, 2002). The interview questions were designed deductively from research on the same topic. Furthermore, the structure of the questions can be led back to the main challenges of the implementation of Big Data derived from literature: namely, data access, data sharing, data analysis, skills and infrastructure. The goal of the interview was also to capture the interviewees' perceptions about the challenges and opportunities of Big Data and its future development within the context of STDs. When creating the interview guide, and all throughout the research process, the interviewer must consider their personal biases and assumptions, as this is critical for the accuracy of the data (J. Creswell, 2003; Magaldi & Berler, 2020). In order to avoid biases, the interview guide was revised by another researcher and leading questions were averted, meaning that questions were designed in a neutral way so as not to subconsciously prompt participants to give a specific reply and thereby restrict their answers.

2.4 Population and sampling procedure

The selection of a sampling method and criteria for recruiting interviewees is critical, as it might have implications on the analysis of the data acquired (Mehmetoglu, 2004). Purposive sampling, snowball sampling, quota sampling, and convenience sampling are some examples of nonprobability samples used by researchers (DeCarlo, 2018). For the purpose of this research, the interview partners were selected purposively to cover a variety of crucial stakeholders (e.g., DMO, IT companies and departments, public sector) within Valencia to represent a detailed and diversified perspective on this issue. Purposive sampling is a technique in which participants are chosen based on parameters such as age or job position that the investigator has established to be relevant to answering the research question (Given & Saumure, 2008; Saunders et al., 2019). In this case, it entails finding and choosing individuals who are particularly educated about or experienced in a relevant topic (J. W. Creswell & Plano Clark, 2011; Saunders et al., 2019). Some criteria that were employed in the recruitment process included: 1) the respondent must be located in Valencia, 2) the interviewee must work in or be engaged with tourism, and 3) the informant must be familiar with Big Data. As a result, the interviewees were chosen from the tourism sector and based on their knowledge. Choosing persons from the population who can provide as much in-depth insight into the issue as possible constitutes an exhaustive sampling unit (Johannessen et al., 2020). However, since recruiting people for the in-depth interviews proved to be difficult, snowball sampling was used to obtain additional respondents afterwards. This strategy entails a small group of individuals referring to other potential interviewees they

personally know (Easterby-Smith et al., 2015). The decision followed the researcher's various attempts to contact knowledgeable tourism professionals through email, LinkedIn and Facebook but not receiving enough responses. As a result, a combination of purposive and snowball sampling was used to identify respondents in the end. Furthermore, the ultimate sample size was determined using the concept of empirical saturation. There was already an indication of saturation after the 6th interview indicating that further data collection would not provide further insights or new thematic concepts (Charmaz, 2006).

Table 2: Overview of interviewees

Sector	**Position**	**Name**
Public sector		
SEGITTUR	Director of Tourism Research, Development and Innovation	Carlos Romero Dexeus
General Fundació Visit Valencia (DMO)	Subdirector	Joan Carles Cambrils i Camarena
Turisme Comunitat Valenciana	Head of Strategic Planning	Francisco José Garcia Martinez
Private sector		
Bioparc Valencia (zoo)	Deputy CEO	Daniel Pons Mallol
Hotel Business Association of Benidorm, Costa Blanca and the Valencian Community (HOSBEC)	Technician in Sustainability, Innovation and Digitization Projects	Jorge Ferrándiz
Fundación Valenciaport	Mobility, Cruises & City-Port Director	Carolina Navarro Correcher
	Innovation & Port Cluster Development	Joan Meseguer Llopis
ForwardKeys	Head of Marketing	Michelle Wols

Source: Own work.

Due to the limitations of scope and time, the sampling size consists of eight informants. These experts and managers come from the public and private tourism and technology sectors in Spain. The interviewees were purposefully chosen since they are all essential stakeholders with expertise and direct engagement in the adaptation of Big Data in Valencia. The goal was to concentrate on those who have a unique or crucial viewpoint on the issue under investigation. Interviews with public-sector tourism professionals provided the researcher with a deeper understanding of the destination management's initiatives and future goals. Meanwhile, the interviews with tourism businesses revealed their perspectives on the use of Big Data in tourism, as well as the potential they see and the challenges they

face. The respondents, as indicated in Table 2 above, came from a variety of backgrounds within the tourist industry. Governments, tourism associations, and businesses were among the participants in order to ensure a holistic picture. Nonetheless, the small sample size does not allow one to derive generalizations from the findings for neither the whole tourism industry in other destinations nor within Valencia or any of the sub-sectors such as Food & Beverage or the accommodation sector. This is also related to the constructivist stance, which acknowledges the existence of multiple inter-subjective realities and therefore does not allow for objective generalizations.

Apart from the DMO of Valencia, also the Turisme Communitat Valenciana, which is the public entity of the Generalitat of Valencia and responsible for the promotion and execution of the tourism policy (Generalitat Valenciana, n.d.), forms part of the public sector respondents, next to the Sociedad Mercantil Estatal para la Gestión de la Innovación y las Tecnologías Turísticas (SEGITTUR). SEGITTUR, which is part of the Ministry of Industry, Trade, and Tourism, is the state-owned trading company in charge of managing and advocating innovation and technologies in the Spanish tourism industry, both in the public and private sector (SEGITTUR, n.d.). In particular, one of their lines of action includes the promotion of smart destinations. As for the private sector, the respondents varied from tourism businesses such as the zoo Bioparc Valencia, over associations like HOSBEC and Valenciaport, to data analysis companies like ForwardKeys. ForwardKeys is a travel intelligence company that developed a B2B database of air travel bookings, which includes detailed information on traveler profiles, their behavior and preferences at the destination, travel patterns, demand, and the impact of events on travel (ForwardKeys, n.d.-a). The company's purpose is to help companies that rely on international travelers, such as tourism organizations, hotels, and retailers, make better strategic decisions by analyzing data from tickets purchased through travel agents and directly from airlines. HOSBEC is a Valencian Community hotel business organisation that provides information, advice, and negotiation representation to more than 300 members (HOSBEC, n.d.). The Fundación Valenciaport is another association that provides services to the port and logistics cluster through applied research, innovation, and training (Fundación Valenciaport, n.d.). It promotes innovation to strengthen the competitiveness of enterprises by delivering specialized training for continual human capital improvement, creating market intelligence services, and giving support.

2.5 Data analysis

This research makes use of qualitative and inductive content and thematic analysis to group the responses of the interviewees into themes (Corbin & Strauss, 2008; Miles & Huberman, 1994). A theme contains essential information that aids in understanding and explaining the subject of the study (Braun & Clarke, 2006). It is a qualitative exploration of the viewpoints of different stakeholders like experts with knowledge regarding the previous, current and future barriers and opportunities of Valencia as a smart destination using Big Data. For the

sake of data analysis, the previously transcribed digital recordings of the interviews were coded using the software NVivo. NVivo is a data management and searching application that allows academic researchers to prove the study's integrity, consistency, and hence trustworthiness (Smyth, 2008). Therefore, the transcriptions were read multiple times and then manually coded line-by-line by identifying themes. Content analysis can be categorized using one of two methods: pre-defined categories or categories derived from the text itself (W. Gibson & Brown, 2009; Stepchenkova, 2012). Following an inductive analysis, ten broader themes were defined in this study to examine the benefits and pitfalls of Big Data in the context of STDs. These preliminary codes were progressively sorted into wider categories in order to find hidden trends (Miles & Huberman, 1994) and relationships in the form of commonalities or contradictions between the various concepts across all cases (Yin, 2009). Codes were repeatedly revised via merging and deletion until finalized themes were established, resulting in increased analytical reliability. It should be noted that certain codes fall into more than one category. Nonetheless, when employing inductive coding, the researcher's prior understanding of the subject will almost certainly impact decision-making which one has to be aware of (Harding, 2015). The researcher of this study comes from an international business background with a specialisation in tourism management and an interest in data analysis but no prior experience in the latter field. This reflective account aims to provide context so that the reader of this study can determine whether the findings and analysis are subjective or objective. Overall, the methodological approaches that have been previously elaborated, the documented research process, and the reflective account all contribute to ensure the reliability, applicability, and credibility of this study.

3 CASE STUDY: VALENCIA, SPAIN

Tourism is a key industry for the Spanish economy, accounting for 12.7 percent of all jobs and 12.4 percent of Spain's GDP in 2019 prior pandemic (Instituto Nacional de Estadística, 2022b). Due to the outbreak of the Coronavirus, which halted both domestic and foreign travel, the latter figure decreased to just 5.5 percent in 2020 (Instituto Nacional de Estadística, 2022b; Nicola et al., 2020). In 2021, Spain welcomed 31.1 million international tourists, up 64.4 percent from the previous year but still 62.7 percent less than the year before the pandemic, when 83.5 million tourists arrived from abroad (Instituto Nacional de Estadística, 2022c), excluding same-day visits and transit travelers. In 2020, the main markets of origin were France (18.7%), Germany (16.7%), and the United Kingdom (13.8%) (ibid.). As for domestic travelers, 135.7 million Spaniards explored their home country in 2021, which is around 22% less than in 2019 (Instituto Nacional de Estadística, n.d.).

When it comes to learning from best practices in a Western context, countries like Spain may be a good place to start. The tourism policy established in Spain's National and Integral Tourism Plan 2012-2015 (Secretaría de Estado de Turismo, 2012) reflects the country's interest in the STD approach, with smart destinations listed as one of the priority lines of action coordinated by SEGITTUR (Celdrán Bernabeu et al., 2016; SEGITTUR, n.d.).

Furthermore, the establishment of a Spanish Network of Smart Cities (Red Española de Ciudades Inteligentes, n.d.) and a Tourist Destination subcommittee (UNE, n.d.) have aided in the conceptualization of STDs and their operational models (Celdrán Bernabeu et al., 2016). This effort culminated in the production of two standards regarding the management system as well as indicators and tools of Smart Tourism Destinations, outlining the requirements for developing a STD governance structure that incorporates the application of technology, innovation, accessibility, and sustainability (AENOR, n.d.).

Also the autonomous communities in Spain have begun to incorporate the STD model into their regional tourism policies, thereby consolidating the efforts advocated at the national level (Celdrán Bernabeu et al., 2016). With the introduction of the "Smart Tourism Destinations of the Valencian Community" initiative in 2014, the Valencian Community stands out as a pioneering destination when it comes to addressing the strategic and operational implications of the STD model (Celdrán Bernabeu et al., 2016; INVATTUR, n.d.). This project focuses on developing knowledge about tourism, with Big Data being one of the most important tools (Celdrán Bernabeu et al., 2016). The Comunitat Valenciana is compromised of three subregions, namely Castellón, Alicante and Valencia (cf. Figure 8; CIS, n.d.). The metropolitan area of Valencia can then again be subdivided into four further regions, with the city of Valencia being the main region (Ajuntament de València & Oficina d'Estadística, 2018). Here, the city of Valencia sticks out as it has won the 2022 European Capital of Smart Tourism Awards and has already implemented numerous smart projects, which will be elaborated upon later. In 2020, it was home to 801,545 inhabitants in the city and 1,565,701 residents in the metropolitan area (Visit Valencia, 2020), making it the third-largest city in Spain. As for tourism, in 2021, it was the fifth most visited autonomous community by foreign travelers in Spain with a total of 4,019,625 international tourist arrivals, accounting for 12.9% of all international arrivals to Spain (Instituto Nacional de Estadística, 2022c). In 2017, the tourism sector accounted for 14.6% of the GDP generated by the Comunitat Valenciana and provided 15.1% of the total employment in that region (Exceltur, 2018). Due to the city being located on the Eastern coast of Spain facing the Mediterranean Sea, visitors and residents alike can enjoy the good weather at one of the beaches along the seven kilometers of coast, which is one of its main tourism assets (Visit Valencia, n.d.). One of the most well-known attractions is the Ciutat de les Arts i les Ciències which is a cultural and architectural museum complex. Other attractions include a total of 48 museums, 22 theaters, and a variety of natural spaces (Visit Valencia, 2020).

The impacts of the COVID-19 pandemic on tourism were also noticeable in Valencia with merely 1,703,493 overnight stays (-68.3%) being recorded in 2020, following many years of growth across the tourism sector, with a somewhat lower drop in domestic demand (-49.9%) than foreign demand (-77.2%) (Visit Valencia, 2020). Valencia's demand thus declined marginally less than other major Spanish cities and the national average (-72.9%), according to data from the National Statistics Institute (INE). This was also reflected in a

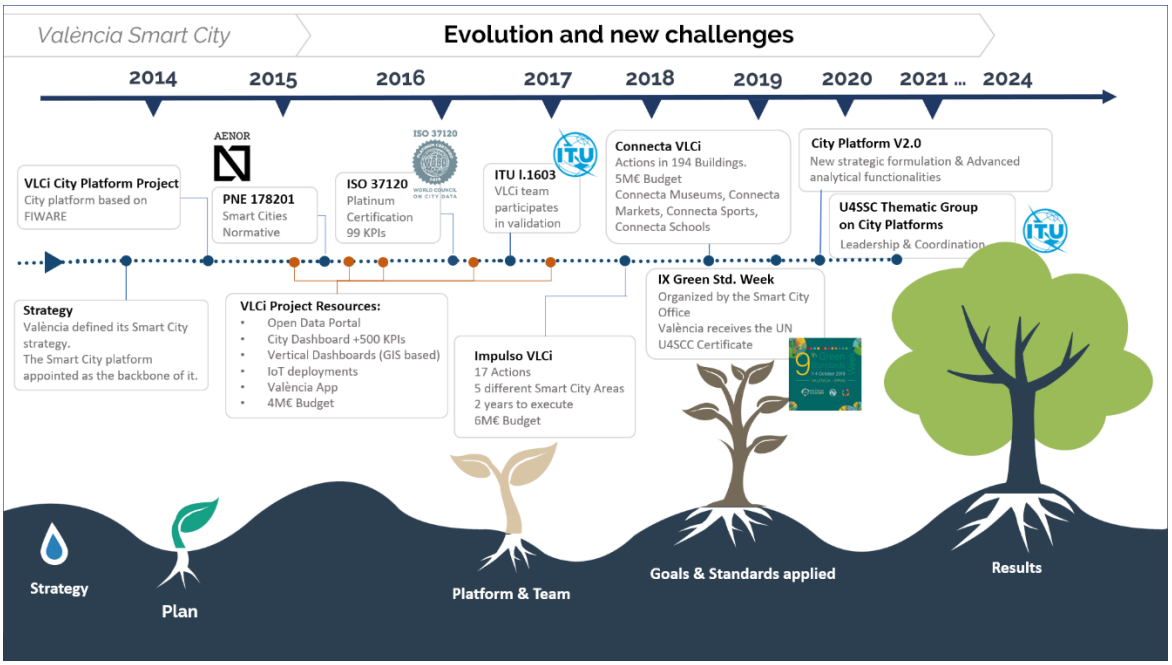
decline in both RevPAR of €31.7 (-52.6%) and the average rate per hotel room of €71.9 (-16%) (Visit Valencia, 2020).

Figure 8: Map of Comunitat Valenciana

Source: CIS (n.d.).

Valencia's evolution to a smart city, and later to a Smart Tourism Destination, already started back in 2014 with the launch of the VLCi platform, an open data and transparency platform (cf. Figure 9; Smart City Office, n.d.). It is a sophisticated computer and storage system that allows local government officials to centralize a significant volume of data from the Valencia City Council systems as well as other systems and devices in the city (Smart City Office, n.d.-g). Data on public transportation, city bike use, and tourist buying patterns using the VLCi IMPULSE card is captured from sensors located across the city. These data are gathered, analyzed with Big Data methods, and then made available to local tourist firms and start-ups for free (European Commission, 2020a). It is also used to create dashboards to aid municipal service administrators in making quick decisions, as well as to generate reports to assist businesses in improving their tourism offerings (European Commission, 2020a; Smart City Office, n.d.-g). Open urban data platforms, which link and interpret all data available, are at the center of the digital transformation of cities (Living in EU, n.d.). Data may then be shared with city departments or third-party companies to give citizens and tourists better seamlessly integrated mobile experiences (ibid.). Communities may customize open urban data platforms to meet their specific needs, exchange data with other parties, and hence deliver improved online services to their residents at reduced costs.

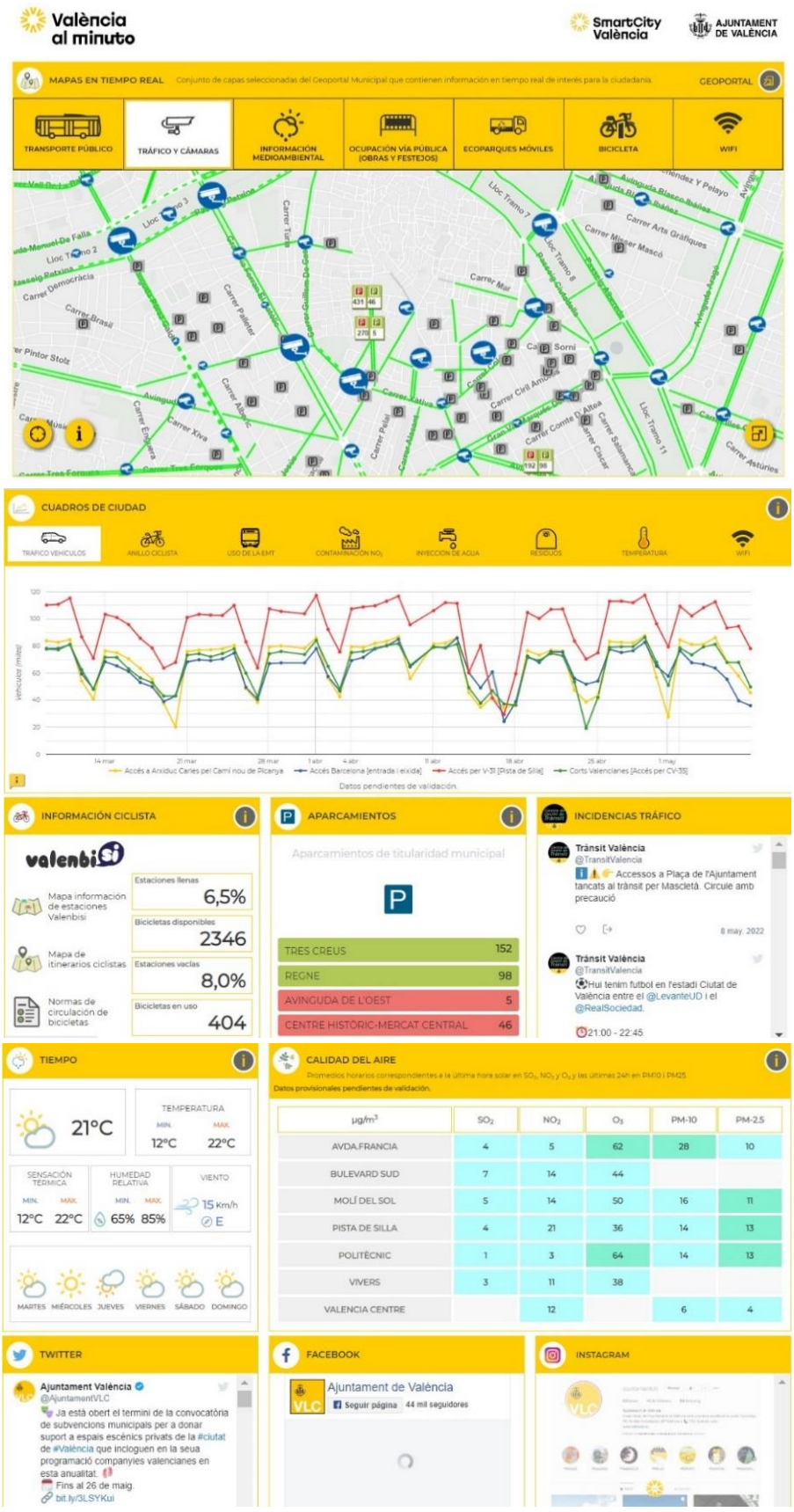
Figure 9: Evolution of Valencia as a smart city



Source: Smart City Office (2021).

In the following year, the Key Performance Indicators (KPI) dashboard was introduced next to the integration of ITC services and sustainable mobility (Smart City Office, n.d.). In 2016, Valencia received the ISO 37120 certification on the platinum level, which describes the indicators to measure the performance of city services and quality of life in sustainable cities and communities (ISO, n.d.). In the same year, another interactive dashboard was launched called Valencia Al Minut, which provides residents and Town Council staff with real-time information on the state of the city like weather information, air quality, or traffic (Oficina de Ciudad Inteligente, n.d.; Smart City Office, n.d.-c, n.d.-f; cf. Figure 10). EMT, a local bus company, was also incorporated to offer real-time information on its services, such as the flow of city buses (Smart City Office, n.d.). The lighting service and public parking were only added in the subsequent year. The local government decided to create the Smart City Office in February 2018 as a separate municipal service with the competencies of advising, managing, monitoring and educating (Smart City Office, n.d.). At the same time, the Impulso VLCi project started which lasted until 2021 and aimed to create a more participatory and inclusive city that improves the quality of life and well-being of Valencia's residents and tourists. Some of the objectives included the extension of the Smart City platform with information regarding the coordination of public road works, gardening services and maintenance of infrastructure, next to the improvement of the GIS platform (IDOM, n.d.). The GIS Portal enables citizens to access a wide range of Valencia City Council's resources for free (Smart City Office, n.d.-b). Displayed on a map, it provides geographical data on topics such as economics, transportation, and social services (Ajuntament de València, n.d.). Employees of the City Services utilize it as well for administrative tasks such as urban planning, property taxation, and infrastructure maintenance (Smart City Office, n.d.-b).

Figure 10: Valencia Al Minut dashboard



Source: Oficina de Ciudad Inteligente (n.d.).

Later in 2019, the Connecta VLCi project, which provides a more effective administration and integration of 194 buildings and municipal facilities with the VLCi platform, was announced (Smart City Office, n.d.-c, n.d.-a). Internal and exterior environmental sensors were installed in many civic places, such as sports facilities, museums, markets, and schools, allowing for management of air and light pollution, temperature, humidity, energy and water usage (Smart City Office, n.d.-a). With the outbreak of the Coronavirus in 2020, the Valencian government released a COVID-19 city dashboard and a website for citizens with information on the spread of the virus and measures (Smart City Office, n.d.). During the same year, Valencia joined the directive of the Spanish Smart Cities network and became a leader in The United for Smart Sustainable Cities (U4SSC) working group on IT solutions for integrated city management and COVID response (Smart City Office, n.d.). The U4SSC is a United Nations program organized by the International Telecommunication Union that provides a global platform for information sharing and partnership development to assist cities and communities in reaching the Sustainable Development Goals (ITU, n.d.). Ultimately, within the scope of the WiFi4EU program, the Smart City Office gradually deployed new high-speed internet access points in public spaces in 2021, thus providing locals and tourists with high-quality internet access (Smart City Office, n.d.-e) whilst simultaneously enabling the government to gain a better understanding of tourist flows and behavior. Thanks to all this work, Valencia scored third in a study of 62 Spanish smart cities in 2016 (Baucells et al., 2016).

In terms of smart tourism, Valencia's Smart City Strategic Plan includes 17 initiatives that, amongst others, also involve long-term tourism plans, all geared toward transforming Valencia into a STD (European Commission, 2020a). As mentioned previously, the Council built the VLCi open data platform to help with this undertaking (European Commission, 2020a). Furthermore, the Tourism Intelligence System (SIT) dashboard was created by Visit Valencia to assist tourism providers and operators in understanding major traveler trends in a straightforward manner (ForwardKeys, n.d.-b; cf. Figure 11). Everything a tourism professional needs, from air traffic and travel demand to hotel occupancy and pricing, is provided in just one dashboard where the data can be simply filtered and compared (ibid.).

Figure 11: Visit Valencia's Tourism Intelligence System

Source: Visit Valencia (2022).

In addition, the Valencian Institute of Tourist Technologies (INVATTUR) has been working on the Smart Tourist Destinations of the Valencian Community project since 2014, with the objective of defining the STD concept (INVATTUR, n.d.). One of the measures included the placement of integrated NFC terminals at roughly 300 locations across Valencia, including museums, landmarks, and tourism enterprises (Smart City Office, n.d.-d). These readers are assisted by smart visitor center offices that are operating 24/7, enabling tourists to obtain information and purchase touristic commodities. When finalized, this project will be combined with the digitalization of the MOBILIS card, a public transport card, to equip Valencia with a vast network providing integrated data available for tourist consumption, sale, and consultation (Smart City Office, n.d.-d). This infrastructure will be of benefit to tourists, resulting in higher customer satisfaction and lesser effects of tourism on the city, leading to more sustainable tourism (Smart City Office, n.d.-d). Another example is the AlterEco Valencia app, in which the collected data is used to redirect tourists exploring the city to new routes and thereby reduce the concentration of tourists in hotspots (European Commission, 2020a). The initiative thereby aims to strike a balance between the need for tourism-related economic development and sustainability whilst mitigating mass tourism (European Commission, 2020a; IVE, n.d.). The project's results will assist policymakers in making better-informed decisions regarding tourism governance while also improving the coordination of actions between private and public stakeholders at both the local and regional levels (IVE, n.d.). The project also aims to set Valencia apart by fostering local customs in a gamified way (European Commission, 2020a). Valencia is therefore an interesting example to look at what successful smart destinations have learned from their experience and how the application of Big Data has affected their development.

4 FINDINGS AND DISCUSSION

In this chapter, the ten themes that emerged from the data collection are investigated. The interviewees' responses regarding data sources, analysis, and the scope of application of Big Data are examined first, followed by their perceptions of the opportunities and challenges of Big Data and the themes' interrelations. These findings will be first be presented and then discussed in light of current research to assess parallels and discrepancies, the relevance of the findings, and, finally, to answer the research question.

4.1 Assessment of the current state of Big Data utilization in Valencia

Before going into the opportunities and challenges of leveraging Big Data mentioned by the interviewees, the status quo of the participating entities' usage of data sources, their current approach to data analysis, and the scope of application of Big Data are being assessed. This will enable the reader to understand the background and context of the interviewees' responses better.

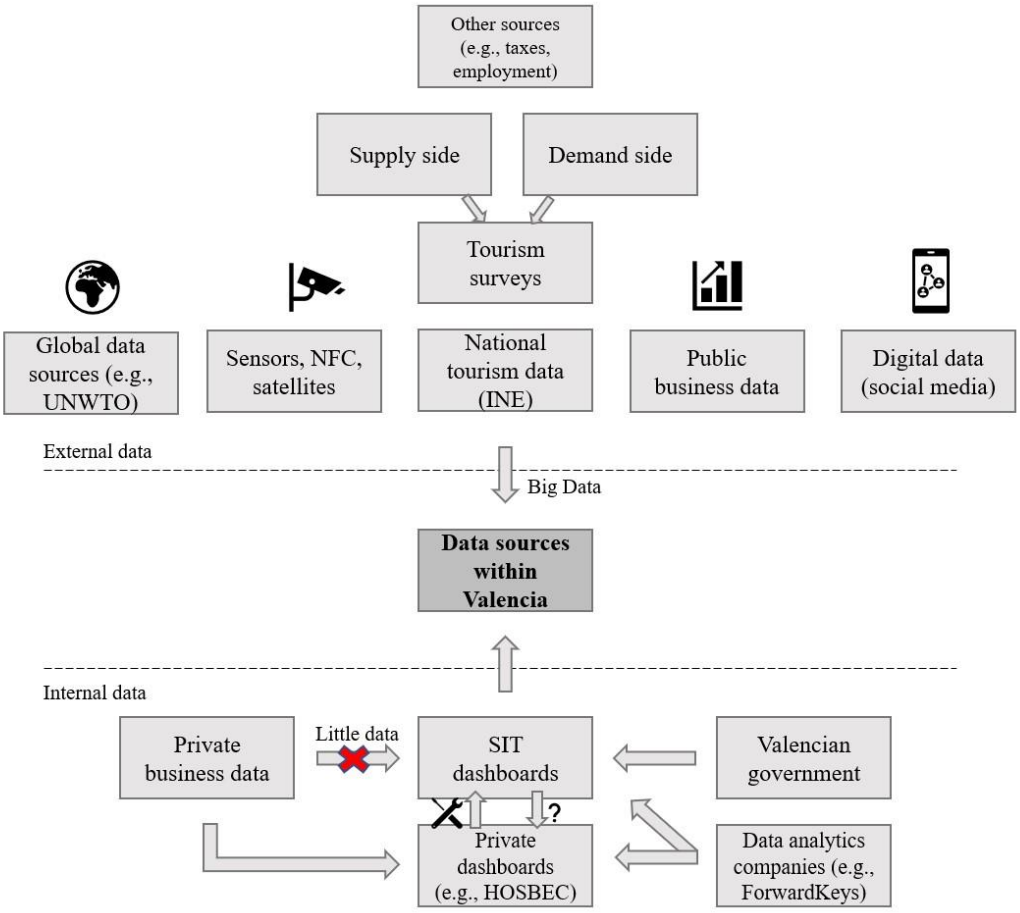
4.1.1 Data collection and sharing

The access to and collection of data lays the foundation of Big Data and data analytics. Therefore, it is important to examine where tourism actors currently get their information from, whether they receive or share those insights from/with other tourism stakeholders or platforms, particularly the VLCi platform, and lastly, what kind of data they assess to be of importance for the future development.

4.1.1.1 Data sources

In the scope of this research, the data sources were divided into local and national as well as global data. Figure 12 shows a summary of all data sources mentioned by the participants of this research and will be explained in more detail in the following.

Figure 12: Overview of data sources in Valencia



Source: Own work (Adapted from Ivars Baidal, 2014).

As for local and national data, the participants of this research are collecting **industry-specific data** according to their field of operation. This means that for example ForwardKeys, which originated as a data analytics company solely concentrated on aviation, mainly collects three data sets: namely, airline tickets booked, flight schedules and flight searches. The zoo in Valencia uses internal data collected from ticket purchases regarding the country of origin, size of groups and how many days in advance tickets are purchased as well as determines customer profiles from its website visitors. The business also has information on visitor expenditures on tickets, in restaurants, and stores. Similarly, HOSBEC primarily works with the occupancy rates of its 300 member hotels in the four markets of Benidorm, Costa Blanca, Valencia and Castellón. Previously, the weekly hotel occupancy data was gathered automatically, but since the pandemic, it has been done manually through an online survey. Through Biontrend, however, which is an analytical tool for hotels and one of the most recent Big Data projects at HOSBEC, the association is connected to the hotels' property management system (PMS), which contains all of the information on bookings and reservations. As a result, HOSBEC receives automated real-time data on occupancy rates, cancellations, revenue, nationalities, booking platforms, and more without the hotels having to do any additional work. This also ensures that the data is

100 percent accurate since the data is not manually collected, which is often subject to errors. Nevertheless, because the integration of a hotel's PMS with the platform is a difficult process, not all members have yet joined Biontrend. In the future, HOSBEC plans to introduce even further data sources, such as flight bookings and online flight searches, as they are trying to develop a predictive model that requires as many sources as possible to determine which ones are useful and whether they are related to the final occupancy rate.

One observation that was made is that some of the interviewed tourism actors, namely Visit Valencia and SEGITTUR, do not gather any data themselves nor have technologies to extract data but purchase it from **external third parties** instead. For example, the Big Data supplier ReviewPro provides information on online evaluations and experiences of visitors who traveled to Spain, which SEGITTUR then purchases and makes available to tourist enterprises. Also Visit Valencia does not collect all the information published on its three SIT dashboards related to accommodation, cruises and air traffic; it is outsourced and supplied by other bodies such as the city hall, the Fundación Valenciaport or the INE. The data provided by the INE is publicly available and used by institutes such as SEGITTUR, Fundación Valenciaport and HOSBEC to receive insights on, for example, the public accommodation sector. Still, those organizations sometimes must acquire even further data for a local statistically representative sample since the INE merely collects statistics at the national and regional levels. The tourism data of the statistical institute is gathered through various surveys and other sources, which are summarized in Table 3. There are four supply-side tourism surveys collecting data from accommodation providers, and three demand-side surveys asking for information from travellers directly. Next to those official statistics, there are also data provided by other authorities such as the Ministry of Labor regarding tourism employment or the Tourism Satellite Account. However, some entities, like the tax agency, only provide limited access to their data. The data from the INE is employed by HOSBEC for their bigger reports, where their internally collected data is combined with public data on overnights, the capacity of hotels, and the number of operating establishments, employees and tourists. Similarly, the Fundación Valenciaport uses data from the INE to see how many people stay overnight in Valencia to obtain certain KPIs. SEGITTUR additionally receives data from AENA, Spain's national airport authority, which discloses information regarding all aircraft arriving at a certain airport, broken down by day and time, airline, kind of plane and number of passengers. Further, AENA has information on slot bookings, which are 15 minutes time slots for an airplane to arrive or depart, four to six months in advance and are useful indicators for anticipating prospective arrivals.

Table 3: National tourism data sources of the INE

Type of source	**Description**
Supply-side surveys	
Hotels	• types of accommodation, number of nights, prices, occupation and distribution channels
Camping sites	
Vacation homes and houses	
Rural accommodation	
Demand-side surveys	
EGATUR (Tourism expenditure survey)	• expenditures, places visited, activities, means of transport, preferences and type of organization of the trip (individual versus packaged tour) • collected on Spain's road, harbour, and airport borders • comes from both foreign and Spanish travellers
FRONTUR (Statistics on tourist movements at the border)	• collected on Spain's borders • similar to a census as it is used to estimate the total monthly amount of visitors arriving to Spain through different means of transportation • done through the combination of information from airport authorities or harbour fronts on the number of vehicles, planes, and their respective occupation with the data from manual counting on road border crossing points on sample days throughout the year
Household survey	• around 10,000 Spanish households are asked about their domestic travel patterns, destinations, modes of transportation and expenditure over a period of time
Other data sources	
Tourism Satellite Account	• national accounts statistical tool to determine the impact of tourism on the Spanish economy • lists all of the tourism characteristic activities
Ministry of Labour	• tourism employment
DERVE (National register of companies)	• enables deriving companies directly related to tourism, such as passenger transportation companies, lodging, restaurants or travel agencies
Tax agency	• allows the quantification of the economic impact of tourism and thereby illustrates the importance of tourism in Spain

Source: Own work.

Apart from these traditional sources, further sources include **digital data**, such as social networks, reviews and online travel agencies which offer a plethora of data concerning searches and bookings. In the virtual environment, one must connect with a variety of providers, such as channel managers, telecom firms, financial institutions, as well as banks, and use various technologies, like computational linguistics, to grasp what travellers are thinking, feeling and saying about Spain in multiple languages, as one interviewee emphasized. Some organizations also receive **local data from governmental authorities**. As previously mentioned, Visit Valencia displays resources provided by the city hall's platform called Valencia Smart City on their website Valencia Al Minut (cf. Figure 13). It shows information about the spread of the Coronavirus, certifications, open data on the local government, and the Geoportal, which is a geographic database. With the help of street maps, aspects like noise and atmospheric pollution, urban ecology, and sports, amongst others, are shown in real-time. Also the Fundación Valenciaport has access to real-time sources as it receives raw data on the number of individuals measured by **sensors** installed by the Valencian Regional Council in busy areas. In return, the foundation shares its gathered information about tourists arriving at the port. Further, it receives real-time data from the Port Community system, which connects the port's actors and allows them to share information, such as on upcoming port calls, cruise transfers and excursions, thereby enhancing the communication between stakeholders in the industry.

In terms of **global data**, only a few interviewees use public data provided by organizations like Google Trends or UNWTO whilst most buy data from data analytics companies such as Mabrian or ForwardKeys. These discovered a commercial viability by combining and analysing official and non-official data and then offering their services to destination managers and businesses. ForwardKeys itself is buying its data from IATA, which in turn receives information on ticket sales and flight schedules from travel agencies and airlines. This flight data is used by Visit Valencia within its SIT as the DMO has good relations with ForwardKeys. Nonetheless, not all tourism stakeholders see the need for obtaining data from external suppliers, like the CEO of Valencia's zoo. When asked whether his organization purchases any data from external providers, he responded with "*No, no, not at all. Why would I need that data?*" (D. Pons Mallol, personal communication, March 25th, 2022).

To sum up, one could say that the interviewed tourism actors attain information from a variety of sources on the local, national and global levels. Yet, many stakeholders see the need to complement this external data with internally collected data. Further, the importance of external data sources is assessed differently by the interviewees.

4.1.1.2 Co-sharing data

When it comes to co-sharing internal data with other organizations at the destination level, most respondents have stated that they are not currently doing this or are in the process of doing so, such as the Fundación Valenciaport. To be more specific, some interviewees do

not share their data online because of administrative problems but do share them individually with all administration levels, while others merely share basic information, such as the number of visitors, on a yearly basis with the tourism offices. Moreover, due to the COVID-19 pandemic, information sharing came to a halt as the obtained data was perceived by the interviewees to not be representative. Ferrándiz (personal communication, April 20th, 2022) added that as businesses recover from the pandemic, they slowly begin to understand the value of sharing data with others. This is causing further tourism actors to want to join data-sharing initiatives and platforms. A similar observation was made by Romero Dexeus (personal communication, March 16th, 2022) when working with governmental entities as he stated that while digitalization facilitated access to the Ministries' data via Application Programming Interfaces (APIs), it is still necessary to explain to them how important it is to provide their information to the Ministry of Tourism and for tourism policies.

Most interviewed organizations only make their data accessible to external actors at a charge. This is the case with Fundación Valenciaport, where the members of the Port Community are integrated and share data within their system, as well as ForwardKeys which provides UNWTO with its data and sells its insights to tourism businesses or destinations. Visit Valencia, on the other hand, discloses its data free of charge. Similarly, SEGITTUR makes its data publicly available on its website, where users may view not only the dashboards but also the raw files. One of the next phases of this organization will be the implementation of an API system and enabling local governments to have daily integrated access to their data. As for HOSBEC, future plans include taking its dashboards to the national level. However, since it is confined to the Valencian Community and its members owing to constraints arising from receiving public funds, HOSBEC itself cannot do so and is therefore collaborating with other national organizations to expand. Further, the Biontrend project, which is a collaborative tool where different organizations contribute information that is then aggregated so they can make decisions based on real data rather than their experience or intuition, is expected to become the standard in several sectors, according to the association.

4.1.1.3 VLCi platform

VLCi is a platform that forms part of the SIT of the Visit Valencia foundation which provides information free of charge on its website to not only the around 400 paid members of the foundation but to everyone who is interested. It consists of both the city dashboards and tourism dashboards that are currently not yet integrated. Further, the Fundación Valenciaport is currently in the process of integrating its platform with the VLCi platform to share its data on the location of cruise passengers in real-time. When asking the participants of the study about their knowledge of the VLCi platform, the responses were very mixed. While some criticized that the information provided is the same for both residents and visitors, which could lead to confusion, others, like the CEO of Bioparc, have not even heard of it yet. However, this was probably due to the CEO not living in Valencia himself. Also the Big Data specialist at HOSBEC has never seen the platform as he believes that only governments

can access it. Visit Valencia stated to raise awareness about the SIT by mailing newsletters to the foundation's members and informing external stakeholders in universities or the media. Many businesses, according to Cambrils i Camarena (personal communication, April 1st, 2022), are aware of the platform but do not utilize it or use it infrequently. As indicated by the findings, one could say that overall, it is not intensively used by many of the interviewed tourism actors and that those have merely low awareness. Nonetheless, this research was limited to only seven tourism entities, meaning that this finding cannot be extrapolated to the whole population of tourism actors within Valencia.

One particular problem that Ferrándiz (personal communication, April 20th, 2022) pointed out is that tools like the Valencian Intelligence Platform developed by associations or local governments are often exclusively restricted to their own members or territory; therefore making it hard to expand. This also has to do with the financing, as the Visit Valencia foundation receives most of its financial support from the Valencian city hall, similar to the previously mentioned data platform of HOSBEC. According to Garcia Martinez (personal communication, April 7th, 2022), some functions that are missing on the platform include information related to the promotional offers for tourists and information about security. Lastly, Ferrándiz (personal communication, April 20th, 2022) criticized that the SIT platform does not provide any new information to them as the data mainly comes from the INE which is already publicly available. In his opinion, the only exclusive data is the one about air traffic. Yet, since his organization is already working with Mabrian, which also provides this information, he sees no added value in using the platform.

Despite the aforementioned critique, it seems that after all some interviewees do see the value offered through the establishments of data platforms such as the VLCi platform as Pons Mallol (personal communication, March 25th, 2022) claimed that the platform's approach "*is going in the right direction*" as it contains a large amount of general information on the city. Further, one has to acknowledge that many destinations' information is not as neatly organized on a single website as Valencia's. Tourism data is sometimes found on one website while local data is found on another; nevertheless, the two should be combined to make it easier for visitors to use local data. According to Cambrils i Camarena (personal communication, April 1st, 2022), there are no longer two distinct domains when it comes to administering a city or managing visitors because they are in constant contact and have similar demands such as traffic management, parking, or air and noise pollution. While one lives in a certain city permanently, the other one lives there temporarily. This relates to Gretzel & Koo's (2021) discussion on merging the smart city and smart destination concept into a smart tourism city concept, which might be a viable undertaking for the future.

4.1.1.4 Important data in the future

One belief that was mentioned by several interviewees is that it is not really about having access to a certain type of data in the future but more about making use, coordinating, and

integrating all of the data that already exists. This refers to the translation of data in a uniform way to enable automated data gathering and sharing. This could be done through the previously mentioned integration of data platforms with the hotels' PMSs, thereby facilitating data collection and making data more accessible. For DMOs, one aspect that is currently missing is the determination of visitor profiles as well as the economic, ecological, and social impacts of tourists. This data could then be used to design marketing campaigns targeting more tourists with a low impact. Visit Valencia aims to employ a new survey that will inquire about characteristics, consumer behavior, and level of satisfaction with Valencian services to work with better data about the demand side of the market. It is thus sticking with traditional information sources. As for the Fundación Valenciaport, it would be interested to receive more information from cruises, such as the number of passengers staying on board in a port and how many book private excursions in comparison to the ones offered by the cruise. The zoo of Valencia, on the other hand, would appreciate data regarding weather forecasts and future flight, train, and hotel reservations in order to estimate the demand. Yet again, Pons Mallol (personal communication, March 25th, 2022) claimed that he could attain that information already by simply talking to the sector, thus not requiring Big Data. Lastly, SEGITTUR declared that data on a weekly or daily basis would be desirable to understand both the past and make better forecasts. It was stated that "*[...] we have been spending a lot of resources and time in processing, collecting, analyz[ing] the past, but we still have a lot to do with understanding the future*" (C. Romero Dexeus, personal communication, March 16th, 2022). Thus, the use of predictive data is assessed to be of importance for the future.

4.1.2 Usage of Big Data

Big Data is used for a variety of purposes by the participants of this research. Many destinations and businesses use data to report and explain historical events (ForwardKeys and SEGITTUR), as a means of comparing themselves to one another, for the preparation of reports (HOSBEC and Visit Valencia), or as an overview of indicators and their progression. Big Data, according to Romero Dexeus (personal communication, March 16th, 2022), is about real-time or daily information, about knowing what visitors are talking about online in different languages and gaining insights from such conversations. The tourism association HOSBEC, for example, is releasing Big Data reports in the form of destination analyses, such as the ones for Visit Benidorm. It considers airline capacities, online flight searches, occupancy, reservations, cancellations, and social media sharing activity by major markets. Aside from reports, HOSBEC additionally sends out weekly press notes regarding hotel occupancy to its members and the interested press and, as previously mentioned, is currently developing its Big Data hotel analysis tool Biontrend.

Apart from generating insights, the sale of the needed Big Data infrastructure poses another business opportunity. Visit Valencia spent over 400,000 euros alone on developing their own computer system and now considers putting this system design up for sale. This might be viable as, according to Cambrils i Camarena (personal communication, April 1st, 2022), in comparison to other cities throughout the world, Valencia is one of the most progressive cities with just these simple dashboards and the plethora of resources it has. As previously mentioned, in many places, tourist data is on one website, while local data can be found on another. Ideally, however, it should be interchanged to facilitate the utilization of local data for visitors. The combination of data coming from these areas may also be useful for the assessment of foreign investment opportunities:

"*I tell you an anecdote: an important hotel company from Barcelona called us about two months ago because they are looking for possibilities for investing in Valencia and these businessmen do not have a big staff searching for different cities and to understand how tourism is working in different cities, medium-sized cities in Europe. And [t]hen we send him to analyze just this data where you can have historical data on how tourism has improved during the last three, four years […]. So just with this data, they have a free and good analysis of information. And all of this information is on the same website. That it is quite useful for foreign investors, for local companies to understand how tourism is working. Not all of them are experts in statistics. Not a lot of them have analysts in their staff. They just want to follow some basic data about the city and tourism."* (J. C. Cambrils i Camarena, personal communication, April 1st, 2022)

This quote proves that simply the accumulation of existing data in an understandable way is already an important managerial tool in today's environment. Within the scope of application of Big Data also lies destination management which will be explained further in section 4.2.2. Next to non-touristic initiatives, the Fundación Valenciaport is working on a number of projects that combine Big Data and tourism. HERIT-DATA is one of them, which intends to promote responsible and sustainable tourism management solutions for cultural heritage sites. This is being accomplished by installing sensors at the port and particular locations across Valencia and using Big Data to better monitor and manage cruise passenger flows, their behavior, and their impact in congested sites. The collected data is then stored and visualized in a platform, which displays the real-time crowdedness of locations so that individuals may select whether to visit one site over another, thus preventing overcrowding.

While some, such as SEGITTUR, make extensive use of data and Big Data sources, others, like Bioparc Valencia, do not at all. The CEO of Bioparc explains this with the fact that since the zoo is in a fixed location with limited opening hours, the insights generated through Big Data analysis are expected to have little impact on its operations. He believes the only impact would be the scheduling of activities and the number of personnel needed.

4.1.3 Data analysis

Data analysis can be subdivided into (1) basic analysis, also known as business intelligence, to analyze and learn from past events in order to be reactive (e.g., through reporting, KPIs, dashboards, or scorecards) and (2) advanced analysis to analyze the current situation and to predict future events and trends in order to be proactive (e.g., through data mining, simulations, or predictive modeling) (RapidMiner, n.d.).

4.1.3.1 Basic analysis

All three HOSBEC, SEGITTUR, and Turisme Communitat Valenciana use the basic analysis methods of reporting, KPIs, and dashboards. Additionally, SEGITTUR also employs conventional statistical analysis techniques such as average values and temporal series analysis. The graphics are then created using a variety of tools and systems, including Microsoft Power BI, Excel, Tabloid, and R2. However, not everyone, such as ForwardKeys, makes use of reporting as the company only writes reports for destinations that are new to data analytics and have little expertise. In that situation, they would create a report or give a presentation outlining the trends in the destination, after which they would brainstorm with the destination about how they might utilize this information. Instead of using Big Data-related technologies, Bioparc Valencia leverages KPIs, data mining, and analysis for internal objectives to better understand the business' performance. Here, a basic database technology, SQL, is used which seems to be a common tool among the interviewees (SEGITTUR, Fundación Valenciaport, and Bioparc).

4.1.3.2 Advanced analysis

Predictive modeling and correlation analysis appear to be popular methods of advanced analysis utilized by the participants (Turisme Communitat Valenciana, SEGITTUR, and ForwardKeys). These may be used to forecast tourism arrivals, bookings, and sales, the impact of tourism, or tourist demand by correlating factors like online searches to local weather or airline capacity. Moreover, given that ForwardKeys only has access to approximately 60% of all flight reservations globally, due to the fact that low-cost carriers do not disclose their data to other entities, the firm utilizes forecasting algorithms to estimate the total global travel market. Bioparc, on the contrary, makes no forecasts. Instead, they run simulations to see how different marketing, price, and product portfolio strategies will perform. Contemporary unexpected occurrences such as wars or natural disasters have demonstrated the necessity for simulations. Two respondents (HOSBEC and Fundación Valenciaport) are therefore attempting to construct a prediction model as one of their next steps, either as part of a project or for the organization as a whole.

4.2 Opportunities of Big Data

While the opportunities created through the utilization of Big Data may be endless, the main aspects mentioned by the interviewees were enhanced decision-making, crises management, destination and business management, predictions and detections of trends, as well as the improvement of the customer experience through personalization. These findings back up the benefits of smart destinations utilizing Big Data that have been advocated within literature (Boes, Buhalis, et al., 2015; Buhalis & Amaranggana, 2015; Celdrán Bernabeu, Mazón, Giner, et al., 2016; Gretzel, Sigala, et al., 2015a; Xiang & Fesenmaier, 2017b, among others). Big Data may therefore be utilized for a wide range of applications, which will be explained in more detail hereafter.

4.2.1 Decision-making and crises management

The findings confirm that one of the most significant opportunities of the implementation of Big Data is enhanced decision-making (Celdrán Bernabeu et al., 2016; Fosso Wamba et al., 2015; McAfee & Brynjolfsson, 2012; SAS, n.d.; Verhoef et al., 2016). This is especially crucial as, according to LaValle (2009), one-third of business executives do not trust the information they use to make choices.

"*[…] if they tell me there are going to be very few visitors this year in Valencia, what do I do? Do I invest more to get them all, or because only a few are going to come do I invest less so I minimize my losses. Because publicity is a cost. So if there are going to be 1 million tourists then maybe I do not need to invest anything, because they will all come. It is not clear what is the impact of having that data in your decision-making. […] the fact of people coming is not related to one variable. It is marketing, it is the weather, the number of people in the city, it is the price. It is hundreds of things. And that is where Big Data could help*."

As Pons Mallol (personal communication, March 25th, 2022) noted above, the tourism sector compromises many variables, making decision-making more complex. Therefore, understanding the interrelations between a variety of factors, next to the different interests of various stakeholders, is crucial to be able to **make better business decisions**, as almost all interviewees stated. This could be attained by utilizing centralised information platform to identify demand patterns and then decide on necessary actions or countermeasures. According to one interviewee, Big Data already impacts the many decisions one makes when planning travels, such as choosing to go a certain way to a specific location. The plethora of available information also enables both consumers and companies to compare products, services, and prices and pick the best bargain. Also in terms of foreign investment decisions, the access to information is vital, as previously mentioned. According to half of the interviewed participants, Big Data can additionally be of help in the **optimization of processes**. It could point out what processes can be enhanced or how the number of people involved in those procedures can be reduced. For instance, if a hotel can predict the number

of clients that will arrive thanks to Big Data, it will be able to determine how much personnel will be required at particular times.

Recent events such as the outbreak of the COVID-19 pandemic or the omnipresent threat of climate change have made it evident that crises from the broader environment can appear at any time. According to some interviewees, open Big Data might help **cope with these crises** in many ways. This may be in terms of measuring the impact of events on the tourism sector or comprehending the interconnection of different factors, which may help in anticipating and understanding how consumers are responding to these crises. Therefore, Big Data may assist businesses to better react and adjust to those changes, thus becoming more resilient to any type of external influence or incident. Open data platforms like Our World in Data or the Travel Industry COVID-19 Monitoring from TravelgateX, for example, might be crucial for fast decision-making and responsiveness, as stated by one interviewee. The latter portal shows the national percentage of cancelations for the most relevant markets, which is a significant variable measuring the impact of crises. These findings are aligned with the theory, which found that as a result of the COVID-19 outbreak, Big Data has become highly important in the tourism industry as it may be used in particularly volatile scenarios to detect changes in demand during and after the event (Gallego & Font, 2021). Big Data may also be employed to track people's movements and keep track of constraints such as the amount of people permitted in confined places, which is critical to know in areas like cruise tourism, as one respondent pointed out. HOSBEC utilized Big Data to compare the occupancy rates before and during the pandemic, allowing the government to determine appropriate subsidies for businesses so that they would survive the crisis. For ForwardKeys and HOSBEC, the pandemic was both a blessing and a curse as there were suddenly no travelers due to the closure of borders. However, instead of losing customers, ForwardKeys gained clients because people were curious about what was happening and when tourism would resume. According to HOSBEC, tourism businesses started to understand the value of data analysis and now want access to the association's data after the tourism sector has begun to rebound. This demonstrates that Big Data is critical for decision-making and crisis management.

4.2.2 Destination and business management

Decision-making is also an essential aspect of managing and promoting a destination or business. According to the findings, many destinations and enterprises use data for tactical or strategic choices, such as identifying (foreign) investment opportunities or **marketing** in terms of deciding which audiences to target. This is consistent with what Xiang and Fesenmaier (2017b) as well as Gallego and Font (2021) address in their studies. The findings indicate that destination managers may determine where and when to execute marketing initiatives to attract visitors by evaluating, for example, the performance of airline connections. Another responsibility that destinations have is negotiating strategic flight

connections as those are key elements for urban tourism. Thus, they must be aware of existing routes and flight schedules, next to those of rivals. When paired with flight searches, it is possible to determine the demand and performance of a route. Visit Valencia then uses this data to plan meetings with local travel agents, especially in those areas where few flights are sold, but the DMO does not want to lose the link to the city. If a connection is underperforming, Visit Valencia contacts the local DMO and conducts marketing activities such as publishing information in local publications and asking influencers to visit Valencia. DMOs may also identify what types of visitors are coming, allowing them to plan, e.g., in which languages information should be provided. Ultimately, the extent of the activities taken based on Big Data varies depending on the tourist board's size.

In principle, academia complies with the finding that Big Data can help gain a comprehensive picture of the beneficial and adverse **impacts of tourism** on destinations on a local, national, and worldwide scale whilst also addressing current tourism challenges such as inadequate management, seasonality or overtourism (Dodds & Butler, 2020; Gretzel & Koo, 2021; Milano et al., 2019; WTTC, 2019b). Instead of attracting as many tourists as possible, according to Wols (personal communication, April 8th, 2022), the present goal of destinations is to keep visitors for as long as possible. Further, more tourists with a considerable economic but low social impact ought to be targeted, thereby ensuring that tourism benefits as many people as possible. This represents another aspect in which Big Data portrays an opportunity: **overtourism**. Big Data can help calculate and manage a destination's capacity by forecasting and regulating tourist flows using real-time data from a variety of data sources (Demunter et al., 2017; Shoval et al., 2014; UNWTO, n.d.-a), thereby preventing overtourism (Bertocchi et al., 2020; J. Li et al., 2018). Tourist behavior, for example, might be instantly captured via video surveillance systems, which many sites already have due to the advent of smart tourism (J. Li et al., 2018). Wols (personal communication, April 8th, 2022) stated that many overtourism destinations, such as Barcelona and Amsterdam, are focusing on limiting those tourists who cause the most harm and attracting more prosperous tourists instead. Venice, for example, has been suffering from overtourism for years and recently introduced a booking system and entry fee for day-trippers to manage the inflow of tourists (Buckley, 2022). Using data to identify where people travel so that the destination may better disperse those visitors and monitor if their efforts are effective might also be a solution to manage overtourism. Fundación Valenciaport's HERIT-DATA initiative is one such project. As for the financial impacts, two interviewees proposed that tourist spending behavior might be determined by purchasing data from credit card firms like Visa or MasterCard. Similarly, it was mentioned that Big Data could be a useful instrument to help private corporations maximize their earnings through enhanced operations and decision-making. Whilst the economic impact of tourism should be maximized, the social and ecological impact, such as carbon emissions, water usage, waste production and the erosion of heritage sites, should be minimized, as recommended by half of the interviewees. One of the other numerous difficulties that tourism boards face nowadays is **sustainability**, which data may help with. In the long run,

it is debatable whether travelers will be able to fly around the world in the same manner they do now, according to Wols (personal communication, April 8th, 2022). It appears that this will be restricted, in which case the DMOs will have to determine which tourists to target. Measuring the environmental implications of tourism and then deciding upon which tourists cause the least harm might be one resolution which, however, does not exist yet. These measures, together with indirect indicators like power consumption and waste generation, may be useful for assessing the destination's tourist carrying capacity and the seasonality of tourism, as one interviewee stated. Here, Big Data could help eliminate the pitfalls of current capacity estimating models regarding data availability. Furthermore, the findings indicate that when attempting to tackle climate change, Big Data as a compilation of data from various sources is required if one wants to understand pollution, resource consumption or waste management. These aspects can be measured through sensors, like it is done in the HERIT-DATA project. Nevertheless, it is important to remember that sustainability encompasses not only environmental aspects, as those were mostly focused on by the interviewees, but also economic and societal ones, which are forming part of triple-bottom-line accounting (Gretzel & Koo, 2021). So far, this has been mostly lacking within smart tourism development, so Big Data might be utilized to collect additional economic and social variables that would help the further understanding of the implications of tourism and tourism policy, according to Romero Dexeus (personal communication, March 16th, 2022). Thus, Big Data aids in acquiring a comprehensive image of tourism's impact on a local, national, and global scale which allows for proactive steps and remedies to potential difficulties to be taken. For example, a location may be congested, and by using Big Data, regulations can be applied to alleviate this situation and so safeguard the local heritage.

Overall, the findings have shown that destination managers may have a better understanding of the tourism industry, its impacts, consumer behavior and experience in terms of (dis)satisfaction, and more thanks to Big Data. Further, it can also be applied to develop new tactics and strategic tools as well as to tackle contemporary issues such as sustainability or overtourism. Nonetheless, so far, Big Data has failed to gain traction in DMOs, particularly in the STD sector, since it is currently used mostly by travel companies (Davenport, 2019; INVATTUR & Territorio Creativo, 2015). Yet, according to the Digital Economy and Society Index, which summarizes indications of Europe's digital performance, Spanish enterprises still fall behind in the implementation of Big Data (European Commission, n.d.-c). In 2020, only 9% of all Spanish businesses employed Big Data, compared to a European average of 14%. Besides that, many governments worldwide have begun using Big Data to assist in the management of smart cities. However, none of the respondents mentioned how Big Data might support improving smart city aspects in general, or the residents' quality of life in particular, which would ultimately affect the tourist experience. That might be due to some professionals still regarding smart cities and smart tourism as separate subjects or because it was overlooked in favor of other factors that were considered more critical.

4.2.3 Predictions and detections of trends

Overall the value of Big Data in smart destinations in terms of making predictions, such as projecting tourism demand (Aiden & Michel, 2013; Celdrán Bernabeu et al., 2016; J. Li et al., 2018), and detecting patterns is further acknowledged within the findings. Tourism businesses are reliant on keeping up with the ever-changing customer behavior and demand as they have been found to be modified by current technological advances (Gretzel, Sigala, et al., 2015; SAS, n.d.; Sheng et al., 2017; Xiang & Fesenmaier, 2017a). Big Data, according to Romero Dexeus (personal communication, March 16th, 2022), has tremendous prospects in multifaceted industries, such as tourism, where there are many interrelated variables, impacts, and actors. By combining data from various sources, one can detect trends or find consumer behavior patterns such as which nationalities book how far in advance, where they go, and how long they stay, as half of the respondents pointed out. To capture the characteristics of the complex tourism ecosystem, it is no longer adequate to focus on a single type of big data, like solely UGC or mobile data. Instead, the use of multi-type data is advised as it is done with Big Data. To forecast the number, flow, or impact of travelers, one must link bookings or social media sentiment to more unforeseeable external multifaceted sets of data that drive tourists to make a decision, like media coverage about a destination or weather, the findings indicated. Li et al. (2018) also suggest utilizing cross-domain data for tourism research, including health, insurance, and educational data, as these ancillary variables might impact tourist preferences. These may then be combined to uncover previously unrecognized patterns (Aiden & Michel, 2013; Alaei et al., 2019; W. Fan & Gordon, 2014; Marine-Roig & Anton Clavé, 2015; Song & Liu, 2017), as it is done with the MOBILIS card and NFC readers across Valencia. From the standpoint of businesses and the relationship with their client, one must comprehend and anticipate customers' preferences. The findings indicate that data may help companies figure out how customers react and whether their tastes are changing so businesses can make necessary changes. The Fundación Valenciaport is already working on predictions to improve situational awareness about tourism's impact as part of the HERIT-DATA project, with the goal of using the data it collects to forecast how many cruise passengers will be in certain places in the short- and medium-term and thus predict overcrowding as well as the impact of cruises.

4.2.4 Enhancement of customer experience and personalization

The findings have shown that the detection of patterns and anticipation of demands as well as preferences that the customers may not even be aware of may aid in enhancing client relationships and improving the quality of the visitor experience by tailoring tourist products and services, which is supported by literature (Andrejevic & Burdon, 2014; Egger, 2013; Gajdošík, 2019; Lopez de Avila, 2015; Rotchanakitumnuai, 2017; Swan, 2013). Two of the interviewees indicated that businesses must adapt to these tastes and collaborate with

customers to develop customized offerings with the goal of conversion into sales. Big Data may further help with customer attraction and loyalty programs, as can be inferred from both the findings and literature (Celdrán Bernabeu et al., 2016). Information, for example on consumer spending, can help to improve the quality of the experience of the visit in Valencia by using it to sell tourists new and alternative ways of discovering their destination, such as through the use of new technologies like apps, augmented or virtual reality, as Navarro Correcher and Meseguer Llopis (personal communication, April 22nd, 2022) suggested. The AlterEco Valencia app is one such example. Moreover, information on different types of experiences, tours, transportation, and hotels can be integrated and promoted to facilitate the trip planning process, hence creating added value, which is one of Big Data's 7 Vs.

4.3 Challenges of Big Data

The obtained results suggest that there are numerous hurdles preventing Big Data from being implemented in STDs. The interviewees' opinions demonstrate the necessity to improve data collection, security, processing, infrastructure, skills, and internal organization.

4.3.1 Data collection and access

One area of concern is the very first step in the Big Data process, namely data access and collection. This is a critical stage because data is at the heart of many smart tourism initiatives (Xiang & Fesenmaier, 2017b). Nonetheless, several challenges arise in this regard. First and foremost, many participants in this study work with data from the INE as the interviewees have stated that **official statistics** have shown to be insightful in the past. However, these tourism statistics are primarily based on surveys (Saluveer et al., 2020), which come with certain drawbacks. For example, the interviewees criticized that the data is only published in retrospect on a monthly basis which was claimed to be insufficient in times of unpredictable change. This does not necessarily imply that real-time data is required, although data on a weekly basis might be beneficial in understanding the past and improving forecasting. This critique can also be found in the literature (Gallego & Font, 2021; Mariani et al., 2018), with academics questioning the validity of official tourism data due to data currency, a lack of standardization in statistical estimating procedures and data collection methods (Lam & McKercher, 2013). Many businesses currently lack a standardized data collection approach, which means Big Data can not be fully exploited (Gorini, n.d.). Further, the respondents pointed out that the credibility of tourism surveys is based on the quality of the answers provided by the companies, which are subject to human errors or biases. Big Data, on the other hand, is more reliable because it is based on actual behaviors rather than stated intentions (Song & Liu, 2017). Despite the fact that Big Data may deliver novel, faster, and more accurate statistics (Gandomi & Haider, 2015; Ozer & Cetin, 2021), there is evidence that DMOs and policymakers are resistant to updating their

current statistical methodology with Big Data models (Batista e Silva et al., 2018; Demunter et al., 2017). Thus, statistics represent a potential field of Big Data application. Another point of criticism regarding official statistics that the respondents expressed was that this information is typically not provided for a specific location but instead for larger tourism areas such as states. If tourism actors wish for more precise data, they must purchase it from big operators on the market. As a result, SEGITTUR frequently obtains non-official information in order to comprehend small and local tourism phenomena better. Yet, not all of the participants believe that these new data sources will be able to replace existing ones, as Romero Dexeus (personal communication, March 16th, 2022) doubts: "*It is not a definitive solution. It is not just that "I am going to close all the traditional surveys and I am just going to use the information from the telecom company." Well, maybe this is a mistake.*" This shows that the novel sources of Big Data should rather be considered supplementary sources that may be used when there is a geographic level where information from the national sources is unavailable. Similarly, academics such as Fan et al. (2014) and Kitchin and Lauriault (2015) suggest using Big Data to complement traditional official statistics instead of replacing them, in order to attain a larger amount of valid data. Ozer and Cetin (2021) argue that several public institutions, such as local governments, customs, and statistical offices, compile a variety of information that, when combined with data acquired by the tourism industry, may be relevant in tourism decision-making. Yet, without the integration of such datasets, their use is mostly unreliable and will not lead to a more competitive and smart destination. A significant issue at the individual company level will be to supplement small business data on customer perceptions acquired through traditional approaches such as surveys with the actual behavior data from Big Data derived from the web or other external sources (Mariani et al., 2018). This would enrich the managerial insights of destinations and tourism businesses whilst adding to the variety of sources which is a characteristic of Big Data.

Furthermore, when obtaining external data, one must be aware that raw data is rarely available and recognize that one will not always be able to work with the same data provider. This may result in non-comparable findings as data differs due to the previously mentioned **lack of data processing and definitions standards**. These differences in definitions are hard to identify as one respondent claimed to get access usually only to a smaller part of the data. Additionally, data providers from the telecom or banking sector may see this type of business as a minor component of their overall operations and may not devote efforts to providing accurate data. They believe it is just a matter of placing their available information on the market but lack comprehension of the types of definitions used in tourism. As a result, the interviewees suggest that tourism businesses and destinations should work closely with data providers to ensure that these are acquainted with tourism-specific data and definitions. Data suppliers must learn to tailor the content they sell to the tourist industry's demands. At the same time, destination managers are currently willing to acquire information from such operators without fully comprehending first what exactly they are buying. In the end, both data suppliers and destination administrators tend to want to simplify things, despite the fact

that the solution is never easy. Cambrils i Camarena (personal communication, April 1st, 2022) brought up this example:

"We have contracted [Mabrian] in 2019 and they gave us some information about sentiment […]. They used to get information at that moment from Instagram and Twitter. […] But when you analyze the results there is a problem with the IPs or with the communication company or where you have your cloud or […] servers. So, the sample shows us that there are a lot of people from the United States, and it is not according to the different types of tourists that we get from the national statistics. Or there are many people where they read their Twitter or their Instagram and it shows that they are from Valencia, but from Valencia in Venezuela, because in Venezuela, there is a city called Valencia that has 2 million population living there. […] Or some other type of analysis that has to be cleaned before they use it in the reports. So sometimes they give you information, but it is not [accurate]."

Many vendors in this industry use scrapping systems that analyze rankings, according to Cambrils i Camarena (personal communication, April 1st, 2022). He also stated that it is known that when individuals are pleased, they frequently do not take the time to rank, leaving rankings ineffective. When someone truly despises a service, though, they seek vengeance. **The validity of online textual data** must be questioned in general, as some visitors may submit fake reviews (Miele & Shockley, 2013), or negative ratings may be removed by social media managers, resulting in distorted images (Dekay, 2012). As a result, software programs that analyze sentiment in online discussions and are able to provide the correct understanding of the context are increasingly crucial. Another challenge that can be derived from the results and literature is the exchange of information across sources (Al Nuaimi et al., 2015; Gorini, n.d.; Komninos et al., 2012; Manyika et al., 2011). The findings indicate that the data exchange between regional and local entities continues to be a barrier because certain regions in Spain have their own procedures for collecting regional tourist statistics, which differ from local data collection methods. Further, though plenty of data can be found in most cities, it generally does not always cascade down to local decision-makers and is dispersed among several organizations (Al Nuaimi et al., 2015; Gorini, n.d.). At the moment, SEGITTUR, for example, follows a top-down approach as it gathers data on the national level and then passes it down to the regional and local levels. The implementation of a bottom-up strategy, however, comes with certain problems:

"So, you have to plan always in the case of Spain on those three levels, so national, regional and local. Traditionally we do not use the information they produce at the regional or local level. What we do is to provide them with our data. So usually the flow goes from the national level to the regional or local. But it is very difficult for the local authorities […] to articulate processes by which I could have access on a monthly basis to regional and local data. So for me, if I have to do a study of the tourism evolution of small tourism villages in Spain, it is easier to do it without asking those villages. Directly get in touch with a telecom company or a big bank to get information from them and exploit this information on a local level." (C. Romero Dexeus, personal communication, March 16th, 2022)

This shows that local capacities should be strengthened to enable the collection and evaluation of timely and reliable data. However, this requires financing. Further investment is needed to acquire Big Data from external sources and analyze it. One issue that most of the respondents identified is that large volumes of data are highly beneficial for making choices but are too expensive for small and medium-sized businesses and destinations, like Valencia, to purchase from an external source, thus representing a barrier. This is supported by literature that explains this with the oligarchic pricing dominance of international distribution system suppliers (Tyan et al., 2020), leading to increased expenses for tourism enterprises (Zupan Korže, 2019). Demunter et al. (2017) also criticize the fragmentation of the market for Big Data sources. This is a disadvantage for tourism destinations, as they confront technological and financial hurdles in obtaining data (Gallego & Font, 2021).

4.3.2 Data privacy and security

Another significant aspect that should not be forgotten when exchanging data is security issues and confidentiality. Some firms may be hesitant to provide extensive details about their operations, as is the case with HOSBEC's members. While most hotels accept sharing information on occupancy rates with the association, they have greater difficulty disclosing more sensitive information such as about revenue. When combining data from a variety of providers, trust must be established with both customers and suppliers, as one interviewee suggested. It is essential to explain to clients why the obtained data is needed and how it will be utilized. This might be for business objectives, such as providing targeted offers, or company reporting purposes. Finally, while transmitting sensitive data, one must adhere to GDPR. Otherwise, privacy concerns prevent tourism stakeholders from providing crucial information (J. Li et al., 2018), which smart tourism highly depends on (Gretzel, Reino, et al., 2015). Maintaining anonymity and privacy of personal data is essential with Big Data (Al Nuaimi et al., 2015; I. Hashem et al., 2016; M. A. Khan et al., 2014; Kontogianni & Alepis, 2020; J. Li et al., 2018; Manyika et al., 2011). The signing of nondisclosure agreements, removal of sensitive information (J. Li et al., 2018), or the implementation of Blockchain technologies might be feasible solutions to this issue. Other solutions include using solely anonymized data, like ForwardKeys does, or storing all sensible data locally instead of in the cloud, as is the case with Fundación Valenciaport. Nonetheless, it has to be noted that the theme of data security was not addressed in detail by most of the interviewees. Similar results have been found by Celdrán Bernabeu et al. (2016), where the Valencian tourism professionals assessed the protection of privacy as not an issue that could be a hindrance to the entry of Big Data. In contrast to this, however, the experts within that study establish it as being one of the fundamental problems, just like academia (Gretzel, Reino, et al., 2015). The researchers explained this gap with the lack of knowledge about certain aspects of Big Data, like the legislation concerning the processing of personal data. In the future, it would thus be advisable to educate tourism businesses on these issues.

4.3.3 Data processing and management

Once the data has been collected, the next step in the process of deriving insights and value from Big Data is **data processing**. As a result of ongoing technological improvements and the growth in smart devices in urban areas leading to an increasing complexity of data sources and data formats, data processing has grown more intricate (Al Nuaimi et al., 2015; Kambatla et al., 2014; Mariani et al., 2018; Rodríguez-Molano et al., 2018). Given that Big Data consists of a wide range of data from many sources, the respondents advocated for effective **integration** methods to be developed. All the data must be collected, translated, and lastly disseminated. For example, Cambrils i Camarena (personal communication, April 1st, 2022) suggests that having an aggregated pool of financial institutions integrating all their data of foreign transactions, rather than just one bank providing data that accounts for only a small percentage of total transactions, would provide better insights. However, one respondent pointed out that technical issues may occur when data suppliers are required to provide their data in accordance with their customers' database and Extract, Transform and Load structures. Hence, it should be aimed at making data universally available to all stakeholders using standardized formats and interfaces (Katal et al., 2013; Lee et al., 2013) to avoid this issue. This is also one of the aspects of velocity, which is part of the 7 Vs of Big Data, as the fast-moving flows of Big Data must be gathered, stored, and analyzed in a uniform and intelligent manner (M. Chen et al., 2014; Xiang & Fesenmaier, 2017b). Furthermore, one interviewee reminded to be mindful of data swamps, which occur when a large amount of unmanaged and unorganized data is retained from a variety of sources, all of which serve different objectives but are often aligned, i.e., duplicated. In other circumstances, due to coverage difficulties or because various types of tourists are included, one data set may actually provide a completely different image than another, findings indicate. As a result, one must learn which data set to employ in given situations, as well as how to comprehend this information.

Data management has been identified as another main challenge in both findings and literature (Gretzel, Sigala, et al., 2015; Song & Liu, 2017). Some data tools like the previously mentioned scrapping programs are offered free of charge but fail to distinguish between noise, accidents, and essential data, which has become indispensable. This is where the fourth V, veracity, comes into play which is concerned with the accuracy of the data (Baggio, 2016; IBM, 2014; White, 2012). Data must be kept free of biases and noise (Song & Liu, 2017) which might be introduced through, for example, crises like the pandemic. Several respondents claimed that the data collected during these periods did not have a statistically meaningful representation. One best-case example that avoids these issues is Biontrend, where the connection to the hotel PMS enables automatic real-time access to data, providing HOSBEC with true independence and information of real value. These automated methods are projected to make data collecting activities considerably easier and accessible for everyone in the future. Another area of improvement includes having higher **quality data** in the future, especially on consumer expenditure and customer behavior. Capturing useful

and quality data rather than a large amount of data and then deciding what is useful will be critical, especially because businesses do not have time to analyze plenty of data every day, as brought up by two interviewees. Eventually, poor quality or invalid data has no chance of helping in making the best decisions; instead, it will most likely waste corporate resources (Bose, 2009; Fosso Wamba et al., 2015) and may result in false correlations, potentially leading to the company misinterpreting a business opportunity (White, 2012).

4.3.4 Infrastructure

Infrastructure is a critical enabler for the implementation of Big Data, next to financial investment, human skills and organizational change. The deployment of Big Data requires a wide range of infrastructure in the form of storage and processing capabilities, sensors, data management software, and much more. The **storage** of the massive volumes of data created and gathered every day is an integral part of the infrastructure required for Big Data. The interviewed organizations either utilize solely external cloud servers and services or in combination with their own cloud servers. The findings indicate that almost all data is now stored on cloud servers such as Microsoft Azure, Dropbox, or Google Drive, which is also supported by literature (Katal et al., 2013; Manyika et al., 2011; Mariani et al., 2018). After all, Navarro Correcher and Meseguer Llopis (personal communication, April 22nd, 2022) believe that the tendency is to go toward the cloud so that fewer physical servers are retained on-premises since those must be maintained. As for the visualization of data within dashboards, Power BI is a popular tool among the participants in this study.

However, apart from bigger companies such as hotel or restaurant chains, online travel agencies, and tour operators, according to Romero Dexeus (personal communication, March 16th, 2022), most destinations and tourism organizations lack Big Data **infrastructure**. Yet, in 2022, around 99.8% of Spanish businesses are SMEs (Instituto Nacional de Estadística, 2022a), meaning that those most likely do not currently employ Big Data. While even some of the larger tourism players may not necessarily leverage Big Data at the moment, they do have bigger data repositories of information on their customers already, as well as sophisticated customer relationship, property and content management systems, as mentioned by two interviewees. Only a few have Big Data on a destination level yet, findings suggest. Since those are expensive systems, most destinations employ external IoT infrastructure, management, and data insight suppliers. Some interviewees believe that in the future, Big Data infrastructure will be outsourced entirely because it is cheaper, more convenient, up-to-date, and efficient, and since specialist organizations are typically more experienced and competent for this transformation. Literature also regards this as a viable option that would address challenges of Big Data such as the lack of in-house competencies, especially in SMEs (Ardagna et al., 2016). Valencia stands out in this regard as it has chosen to develop its own smart tourism system instead of **outsourcing** the competencies. This was

done with the assistance of ForwardKeys, which supplies the city with data and the organization's dashboard. When asked about future infrastructure needs, most interviewees responded that they had all the necessary equipment and technologies and that those were up to date. This divergence to the previously mentioned lack of Big Data infrastructure could be caused by several factors. Either the interviewed tourism professionals were not experts in IT and therefore did not know what other infrastructures exist, or the interviewees were thinking in terms of their current capabilities and not the future potential of infrastructure. According to Navarro Correcher and Meseguer Llopis (personal communication, April 22nd, 2022), staying updated on technology is critical and simultaneously a difficulty in IT, which is also once again related to the need for funding.

In order to go forward with the advancement of Big Data, **digitalization** is vital. Academics like Jovicic (2019) and Jasrotia and Gangotia (2018) state that nowadays it is unfeasible to "survive" without the use of digital and smart technologies and the adaptation of business strategies (Celdrán Bernabeu et al., 2016; Gretzel, Sigala, et al., 2015). While the findings show that digitalization has already facilitated the access to data via API and made it easier to specify variables, challenges persist. For starters, one respondent declared that all digital developments must be effectively communicated so that people understand how and why to adopt new technologies. Furthermore, as not everyone has digitalized their data or uses the internet yet, digitization takes time. Nevertheless, this is necessary before data can be accessed and shared automatically and in real-time.

4.3.5 Data analysis skills

Many interviewees agreed that one of the biggest challenges in the implementation of Big Data is related to data analysis skills. As previously mentioned, most interviewees employ basic analysis, such as reporting, rather than advanced analysis, such as predictive analytics. While some interviewees are working on or have already employed a prediction model, others do not see advanced analysis being implemented in the foreseeable future. However, advanced analysis capabilities are necessary to move further with Big Data analysis.

"[T]he data itself does not give you anything; you need to transform it to get your conclusion" (J. Ferrándiz, personal conversation, April 20th, 2022)

This statement of one of the interviewees summarizes perfectly that data itself is not valuable without someone who evaluates and generates insights from it. It also reflects the need for educated tourism personnel who have a **digital mindset** and possess the **abilities** to deal with any kind of technology and device, as well as with digital markets and their tools, such as Python, Power BI, or machine learning. Certain destinations, as Barcelona, have already begun to install a large number of sensors, thereby greatly increasing their capacity of acquired data, as one interviewee explained. However, they still lack the ability to analyse

it. According to Cambrils i Camarena (personal communication, April 1st, 2022), most European cities, including Valencia, cannot afford to hire a large internal analytical staff and instead rely solely on a handful of specialists. Data scientists with technical and analytical skills are needed to clean, organize and manage the quality of the huge amounts of data (McAfee & Brynjolfsson, 2012) in order for the data to provide value to organizations (Al Nuaimi et al., 2015; Fosso Wamba et al., 2015; Lim et al., 2018; Miele & Shockley, 2013). Furthermore, the findings indicate that in some countries, such as Spain, workers tend to have a lifetime job. However, as the requirements and needed skills have changed, such as tourists now requiring a different interaction on online platforms, there is a greater challenge in getting tourism workers educated, willing and capable of working with data. Almost all respondents agree that nowadays the most valuable resource in a tourism destination, particularly in Spain, is someone with a background in data analytics or statistics. This includes understanding how to work with data evidence rather than intuition, how to get insights from data, how to integrate it, and finally, how to visualize it - the seventh V of Big Data (Baggio, 2016; El Alaoui et al., 2017). In 2020, 60% of the population in Spain had at least **basic digital skills**, which is lower than in the rest of the Economic and Monetary Union (Arellano et al., 2021). These observations are backed up by literature that states that many organizations lack the skills needed to extract insights from Big Data (Manyika et al., 2011; Song & Liu, 2017). The tourism sector does not yet draw in highly skilled employees, despite its dependency on ICTs (Gretzel, Sigala, et al., 2015). This is mostly a matter of setting up training programs, according to two interviewees, as the employees' degree of education also has an impact. One respondent observed that currently, most tourism students do not graduate with a thorough understanding of digital tourism. Also academia advocates that academic institutions must invest in Big Data curriculums to generate skilled professionals (Katal et al., 2013) as a response to the lack of Big Data academic programs for tourism management (Mariani et al., 2018). Yet it can be seen that gradually more competent employees are hired in the tourism sector, even though it differs in each company, as noticed by one respondent.

Transforming a company or a destination to be more data-driven seems to involve a major **investment in human resources**. The findings back up academic assertions that IT investments and the development of data analytics capabilities are required (Bertot & Choi, 2013; Frederiksen, 2012; Sheng et al., 2017). If business owners are reluctant to invest in human skills, it is likely that a major part of data-related tasks will be outsourced to data analytics companies specializing in specific fields, such as ForwardKeys or Mabrian. Half of the respondents believe that hiring those services is faster, easier, and maybe even cheaper as they have enhanced investigations, data accuracy, and data cleaning. Others, on the other hand, argue that externalizing technicians actually raises expenses, lowers personalisation, and limits flexibility to modify operations on short notice. According to the literature, the viability of outsourcing depends on the skills in question to be outsourced or developed in-house (Doran et al., 2020). Regardless, the findings suggest that businesses must have at least basic internal skills to comprehend third-party information, how to analyze them, and

what questions to ask. Otherwise, in some economies, a lack of competencies has resulted in ICTs failing to deliver the promised benefits (Morales-Urrutia et al., 2020).

In the future, findings indicate that combining **operational and technical expertise** will be the most effective way to gain a more strategic perspective and enable non-data scientists to comprehend the complex potentials of Big Data for their industry, to which Katal et al. (2013) add interpretive and creative competencies. Likewise, knowing more about the industry might help Big Data scientists provide better insights. Particularly the need for multi- and interdisciplinary collaboration between data scientists and (tourism) management is supported by literature (Fuchs et al., 2014; Mariani et al., 2018). Nonetheless, this is a complex and time-consuming process, and little steps must be performed first, such as migrating day-to-day operations, like administrative systems, to technologies so that the operations' outcomes may be immediately used for Big Data analytics, as is indicated by the results. While academia suggests that small and medium-sized tourism-related companies that cannot afford to hire data scientists may purchase reports produced by data analytics firms (Mariani et al., 2018), Ferrándiz (personal communication, April 20th, 2022) forecasted that all SMEs will have at least one data analyst, who will not only act as a technician but also take on further responsibilities. In the long run, having qualified personnel will save businesses time and money.

4.3.6 Internal change

As with any other type of major transformation, **initiating the change within organizations and industries** poses a barrier. Some interviewees proposed that data dashboards should be created by national or regional authorities, who should purchase data sources on a local, regional, state, and even national level, and then offer it to tourism enterprises. This is also supported by literature, which claims that the government and the public sector are primarily responsible for advancing BDA (Kambatla et al., 2014). However, one interviewee revealed that only few dashboards have been produced by most regional administrations so far, meaning that they lack experience. As for the national authorities, several data projects have previously been proposed by the EU through EUROSTAT and the Spanish government. This year, the Spanish government plans to invest 100 million euros in a Big Data bureau where they will purchase information on, for example, residences, according to Cambrils i Camarena (personal communication, April 1st, 2022). The one-off cost of the government designing a platform that collects information from networks and then presenting it at a local, regional, and national level or the accumulated expenses of individual companies buying information from external data analytics companies is estimated to be the same. Small and medium-sized businesses or destinations that could otherwise not afford to purchase external information would then be able to access this vital data and make decisions based on it, thanks to the platform. This is comparable to what SEGITTUR and the SIT of Valencia are

already doing. However, in this case, the data would come from local and regional sources rather than solely national ones. This relates to the argument that investing in Big Data infrastructure, data access and analysis is essential but not feasible for everyone. The majority of Valencian enterprises polled in Celdrán Bernabeu et al.'s (2016) study claimed that they are still hesitant to invest in Big Data because of the large amount of money necessary, a lack of the right organizational culture, or because they do not have sufficient experts to deal with the challenge. ForwardKeys strives to assist and educate its clients in the usage of data, but it is a complicated process, and if businesses just have one data scientist, they must adapt internally first. Enterprises should be the ones responsible for educating their personnel because, in the end, not every company might need to utilize Big Data, according to Ferrándiz (personal communication, April 20th, 2022).

"I am not a hundred percent sure how we can use it. […] we could use it if you are able to correlate so many different factors but I am not sure how you take what the Big Data tells you, how you present it in your decision process. I think there is still a lot of work to be done and it is very important to understand […] what is the added value and who captures that value meaning. If you want to do a Big Data project, there is a lot of investment to be done by somebody, there is a lot of work to be done in development, in the data analysis, data scientists, engineers, etc, and those people need to be paid. So the question is: how do you generate all that value, how does the person or the entity that invests in that effort, get paid? In order to get paid one has to generate more value because otherwise, nobody will pay him and that is for me a business model problem. I mean I am not sure how somebody is going to make money with this." (D. Pons Mallol, personal communication, March 25th, 2022)

"[…]one of the challenges that we are facing is the changes in processes inside companies. For example, there is a big resistance to change in some areas of companies, especially when the employee does not understand the benefits of Big Data." (J. Ferrándiz, personal conversation, April 20th, 2022)

Understanding the added value, which is one of the seven Vs of Big Data (Baggio, 2016; Celdrán Bernabeu et al., 2016; Pérez, 2015), and how to exploit Big Data is another major difficulty, as presented in the quotes above. The findings have shown that not everyone, such as the zoo, uses Big Data yet since they regard its usefulness as limited. While many tourism firms would like to employ this technology, once they receive the data, the issue of what they want to do with it or the realisation that they already had the data before arises. This finding is not surprising, as various studies have indicated that many cities, governments, and executives typically lack an understanding of Big Data's utility and are uncertain what to do with it (Gorini, n.d.; Manyika et al., 2011). As a result, data insights are not incorporated into action plans (Gorini, n.d.). In certain organizations, a lack of understanding can also lead to strong **resistance to change**, especially when employees are unaware of the benefits of Big Data, as one interviewee claimed. Findings advocate Shah et al.'s (2012) discussion that data knowledge must be funneled down to lower levels since analytical skills are still primarily limited to the expert level. While the companies' executives may

understand the value of Big Data, if that data culture is not passed down to all employees of the firm, some of them may be unwilling to make changes in their work processes, and therefore will not take the actions necessary to initiate the process. It is the top level that must demonstrate the value of Big Data in the company and drive other workers to strive towards it. To gain more acceptance, data must be made understandable and organized in an intelligible manner (Bose, 2009; Gorini, n.d.) which is what HOSBEC has done:

"We like to say that we passed from Big Data to smart data because we now make all of that database that we had comprehensible and we make intuitive reports [because]our members […] do not always have a really high formation about this. So, making it accessible for them and easy to understand for them has been key in the last few years." (J. Ferrándiz, personal communication, April 20th, 2022)

Similarly, the findings suggest that one cannot expect local businesses to make data-based decisions if already DMOs are struggling to do so. Thus, a change in the internal organization is required, which takes time, and Big Data technologies must be implemented in accordance with a holistic strategy. Nonetheless, there might be so-called "attitude-behaviour gaps" that emerge when people say one thing because it is socially desirable or expected from them but then act another way (Eberhart, 2015). Thus, the willingness to employ Big Data must not be merely stated but has to be backed up with actions such as through a change of business strategy as may be inferred from annual reports.

4.4 Key aspects for the future development of Big Data

The implementation of Big Data in Smart Tourism Destinations and in tourism businesses is well on its way as currently the first steps are being taken. While other sectors like logistics are already more digitalized and thus faster in the adoption of new technologies, the tourism sector still lags behind, as claimed by two of the interviewees. **Digitalization** and **automatization** were also some of the key aspects mentioned by the participants of this research that are vital for the future development of Big Data. Manual data collection represents a barrier to the employees' work since it demands more effort and time, which is not always possible, such as when hotels are close to full capacity. This is compounded by the fact that not all data is captured digitally yet, making sharing, processing, and tracking data even more challenging. Once all data has been digitalized, the automation process may allow for a continuous flow of data, as the findings indicate. Therefore, some interviewees recommended investments in the advancement of infrastructure, such as 5G and IoT technologies. Programs like Horizon Europe, for example, further acknowledge the demand for supranational digital transitions. Nonetheless, even if governments or supranational agencies subsidize data access and data infrastructure, it is unclear to what extent they will also promote the development of a broader smart tourism sector (Gretzel, Reino, et al., 2015). Further, literature suggests that residents and tourists who are digitally illiterate, as

well as those who are disadvantaged, may be left behind within smart destinations (Buhalis & Amaranggana, 2013; Komninos et al., 2012). This was not addressed by any of the interviewees, which is reasonable considering the study's emphasis on the tourism supply-side rather than the demand-side. However, because the smart tourism infostructure is complicated and demands knowledge and investment, digital exclusion affects not just individuals but also tourism providers (Gretzel, Reino, et al., 2015). When smart tourism is adopted in a destination, tourism businesses that are already falling behind in digital innovations may then lag even further behind.

It is also important to remember that incorporating technology into tourism destinations is not a remedy (Baggio et al., 2020) and that no "one-size-fits-all" ICT system exists (Buhalis & Amaranggana, 2015). Tourism actors should not be deceived into thinking that technological improvements alone would make tourism destinations smarter (Baggio et al., 2020). Over-expectations might be created, which, if not met, could lead to increased distrust in destination management activities (ibid.). In literature, there are warnings concerning the risk of using technological tools without first revising the procedures of interest and how this might lead to reduced efficiency or economic growth despite significant investments in ICTs (Brynjolfsson, 1993; Hammer, 1990). Following that, it can be argued that destinations should adopt bottom-up strategies that focus on customizing technological platforms to local requirements rather than just deploying them (Neirotti et al., 2014; Robinson, 2012).

Further crucial aspects in the implementation of Big Data are **training** and **guidance**. Almost all interviewees agree that employees ought to be trained in data analysis and should know the relevant tools. Otherwise, Garcia Martinez (personal communication, April 7th, 2022) fears that the tourism sector might be too late to capitalize on this trend with no or only little return of value. According to Pons Mallol (personal communication, March 25th, 2022):

"*[I]f we had to do a Big Data project, we would need for sure external help because we do not have experience in Big Data. We would know how to ask the questions, so what questions we want to know. But somebody else would have to do the analysis for us, that is for sure. […] I do not think that a company like us should have a Big Data specialist, we should probably get training on what Big Data can do for us to try to understand the right questions and the questions [that] are possible to answer. But it should be a third-party consultant or something that is able to consolidate all the data and give you the answer.*"

Thus, at times external assistance will be needed, especially in the case of SMEs who might lack capacities and funds. This goes together with the need for **willingness** and **understanding** of the benefits of Big Data analysis. This must come from everyone involved in the company in order to be encouraged to use it and adjust one's work processes, as mentioned by half of the respondents. Higher-ranking employees, in particular, must recognize that trusting one's gut instinct is insufficient nowadays and should therefore drive internal change in the company. What is more, tourism actors must devise a strategy for who will manage, evaluate, and take decisions based on this information. Periodicity, costs,

sources, responsibilities, and partners (e.g., technological providers) are all further factors to consider. Last but not least, when developing a Big Data platform on the destination level, Visit Valencia encourages other destinations or authorities to listen to the needs of the local businesses as they usually do not require a lot of complex data given the limited amount of time they can spend on data analysis. Tourism companies mostly want to know the expected number of tourist arrivals, place of origin, purchasing power, and behaviour.

The future development of Big Data is perceived differently by the interviewees. While some respondents believe that Big Data adoption is inevitable, and certain researchers even describe it as a panacea (Mayer-Schönberger & Cukier, 2013; McAfee & Brynjolfsson, 2012), others claim that in the current uncertain climate, businesses will prioritize short-term issues over long-term planning. Nonetheless, the same respondent acknowledges that if someone collects a large amount of data about people, prices, or climate and makes it available, tourism companies can cross-check it with their internal data and thereby enhance operations. The responsibility will shift and be outsourced from clients' insights to corporations, according to Garcia Martinez (personal communication, April 7th, 2022). Further, two respondents assert that the adoption of a Big Data approach will likely occur at different speeds. First, the larger businesses and destinations act swiftly as some are already engaging with Big Data, while the remainder, particularly SMEs, will lag behind and take longer to reach parity. Integrating and automating all related processes will facilitate the usage of Big Data for more businesses as it will be more accessible. In return, those companies will begin to see the benefits of working with Big Data sooner and may become more interested in its usage. Furthermore, the way Big Data is employed in businesses is expected to alter, as the findings suggest. Till now, numerous organizations have been using data to mainly analyze what happened in the past, which is the first step in applying data analytics. The next phase in the usage of Big Data, however, will be to use **predictive data**:

"[…] *once people get more acquainted with how to use the data, then, slowly but surely they will start to move into more of predictive analytics where they will actually look ahead and see how they can use the historical data to predict the future or to understand or anticipate the future. That is a slow and probably also very painful process where managers and directors have an important role to push those questions from changing within the organization*." (M. Wols, personal communication, April 8th, 2022)

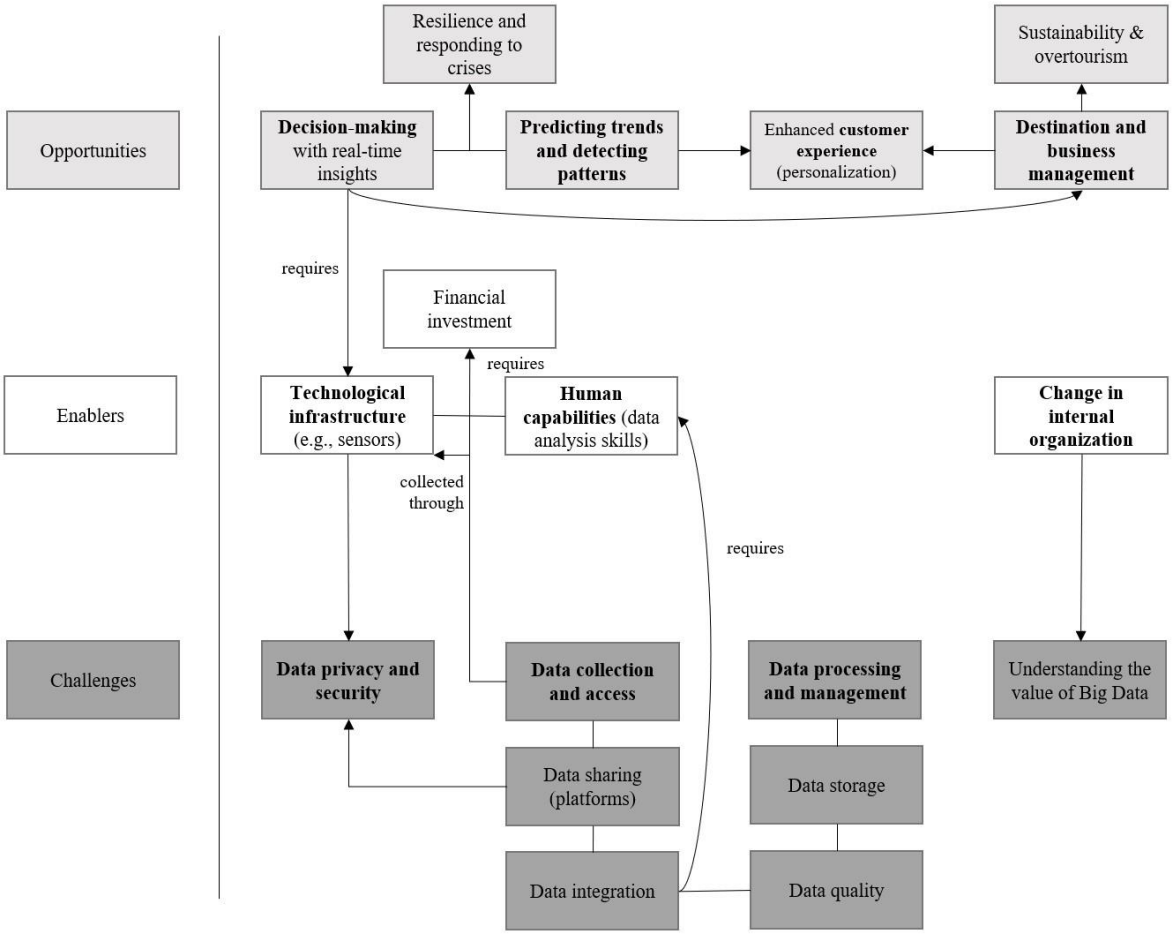
While data analytics companies or destinations may already know the supply side of tourism in terms of how many trains, planes, or cruises will arrive in the future, for the respondents of this research it remains a challenge to estimate the demand, meaning how many tourists will actually buy those available tickets. Finally, these insights will result in a better visitor experience as well as more advanced and personalized services supporting travelers in real-time based on the interconnection of various economic, social, and ecological factors, feedback, and sentiment analysis to create global situational awareness.

4.5 Findings and discussion recap

So, what do these findings imply? This study aimed to discover what opportunities and challenges are faced when utilizing Big Data in Smart Tourism Destinations. By taking the STD of Valencia, Spain as a case study, ten themes have been identified, which are summarized in Figure 13. The key themes were highlighted in bold and divided into three categories: opportunities, enablers, and challenges. However, it must be noted that all enablers represent a challenge at the same time. This depiction also shows that the answer to the research question cannot be simply given within one or two sentences as it is a multifaceted topic in a complex industry with many interrelated variables.

The results were, all in all, consistent with theory, even though past research was mostly conducted in a smart city context. The findings have confirmed that the main opportunities of Big Data lay in assisting STDs and their actors in making better decisions, understanding the past, present and forecasting the future through the detection of patterns and trends, as well as enhanced destination and business management, which all may lead to an overall improved customer experience. Further, the results have supported that Big Data may tackle contemporary challenges that the tourism industry and STDs in particular are facing and attempting to deal with, such as overtourism, sustainability, and crisis management. Lastly, the literature's call for the use of Big Data to overcome the limitations of traditional statistics and data collection methods through gathering more accurate and valid data, next to increased data availability, was advocated within this study. However, in order to take advantage of these prospects, companies must first attain the four enablers for the implementation of Big Data which were identified within the results. It was revealed that destinations and businesses must also deal with internal change, willingness issues, and organisational hurdles, next to the financial, human, and technological barriers that can be found in literature. Challenges occur at almost all steps of the BDA process, from data collection over processing to analysis, the generation of insights and their transformation into actions, as can be derived from both the findings and literature. Nevertheless, when examining the potentials and challenges of a certain phenomenon, several perspectives must be considered. For example, while personalization of products and services is one of the key prospects of Big Data from the perspective of private firms, this is not the case from the point of view of public managers as those are more concerned with tourism policies, impacts, and capacity management.

Figure 13: Summary of key themes and their interrelations



Source: Own work.

The fact that the findings are still in line with literature even though those studies were published already several years ago indicates that the tourism industry has not yet adopted Big Data fully and that tourism actors are facing the same challenges even now. Although it is already well on its way, the tourism industry as a whole is recognized for being a relatively slow adaptor of technological advancements, according to two of the interviewees. Particularly in terms of digitalization, the sector is found to still be lagging. Similar to today, consumers were initially skeptical of the internet's emergence but quickly came to appreciate its new possibilities. Currently, technology and tourism seem to be perceived as two separate fields, but in the future, it would be desirable to merge them as tourism becomes increasingly reliant on technological advancements. Given the ongoing increase in technological solutions employed in practically every sector, basic technical skills, next to operational ones, within tourism will be of utmost importance when going forward. This is the case especially in Spain, as the country's workforce currently lacks basic digital skills. These ought to be incorporated into the curricula of schools and universities so that graduates will possess this expertise. Further on, continuous training will be needed to adapt to the ever-changing work requirements and customer demands. In the end, the value of data is only generated if it is being analyzed and understood by everyone within an entity, as merely

having data is not enough. Taking this action may help employees realize the potentialities of Big Data and lessen their reluctance to initiate the change. It may also be possible to overcome the resistance to change of DMOs, policymakers, and businesses that can be seen in the findings by using Big Data initially in simpler ways, such as by linking weather forecasts to demand predictions, and then gradually evolving. Businesses may also be more inclined to invest in its execution if they are not intimidated by this enormous concept and recognize its value.

Nonetheless, when academia sometimes seems to be idealistic, particularly in the early stages of a new phenomenon, a mismatch between research and the industry might be observed. When it came to the usefulness of Big Data for tourism destinations and enterprises, this was the case, as academia took a more optimistic position, whilst the respondents to this study had a more cautious approach. While they recognize the numerous advantages that this new technology provides, the extent to which it is now used is still limited. The results have demonstrated that some of the tourism industry players frequently view Big Data as a complex and unattainable technology. However, users seem to be unaware that they do not need to fully exploit all its intricate potentialities in order to take advantage of Big Data. Instead, there are uses for it that are much more straightforward. Based on the findings, one could conclude that the present need for data is focused on more basic data such as tourism arrival predictions, nationalities, and weather. Therefore, rather than capturing a large number of complex data, it would be reasonable to save effort by listening to the needs of the industry and collecting more useful and quality data. Companies will hence find it easier to gradually adapt to this change, as they currently lack the resources to analyze massive amounts of data.

After all, tourism destinations and all associated stakeholders will face a multitude of challenges in terms of smartness development. This indicates the need for a "step-by-step" strategy, in which destinations and enterprises must first fully embrace digitalization as well as acquire Big Data infrastructure, expertise, and funding, next to changing their internal organization before realizing the full potential of the technology. However, the internalization of both infrastructural and human capabilities requires funding which these actors, and especially SMEs, might not have. In the near future, many SMEs as well as small- and medium-sized destinations that are not digitally advanced might continue to fall behind with the deployment of Big Data technologies because they do not yet know or comprehend the benefits of its use and lack the enablers of Big Data, necessitating external assistance in order to close the digital gap. Using external data infrastructure and insight providers offers the chance to facilitate the entire data collection and analysis process by requiring destinations and organizations to only have fundamental data analysis abilities in order to comprehend the obtained insights and put them into practice. This is particularly vital in Spain, as the majority of businesses are SMEs. Thus, support from the government in the form of finances, a holistic data strategy and action plan, and through the establishment of a centralised national or regional platform that integrates data from all levels is needed.

This goes together with a need for improved data management and standards in data collection and processing for easier integration and sharing of national, regional, and local data. Ideally, the sharing of data should be free of charge, so it does not represent a barrier to smaller entities whilst adhering to data privacy regulations. This would also enable the data to cascade down to the lower levels as well as entice local data to be collected in a more reliable manner, therefore combining both a top-down and bottom-up approach in data sharing. When it comes to data privacy and smart cities, another mismatch between academia and the findings can be noted as most of the interviewees did not delve further into these topics. This could be attributed to the interview design, as these subjects were not directly addressed, or to the interviewees' decision to concentrate on other, more relevant matters. Further, while in academia the use of digital data and sensors for understanding visitor flow patterns is being praised, some interviewees were somewhat sceptical. They argued that prior to collecting a sizable amount of real-time data through sensors, it is important to understand why that data is needed and what one plans to do with it. In regard to digital data, concerns remain about its validity. This demonstrates once more that industry concerns regarding the deployment of digital solutions persist, in contrast to the somewhat optimistic stance of academia.

In the case of Valencia, it can be noted that there a currently various dashboards, such as those from HOSBEC or the Fundación Valenciaport, apart from the main SIT dashboard of Visit Valencia which are not all integrated yet. Also the sharing of "little" business data with the data platform remains limited as of today, according to the findings of this research. However, the provision of additional more varied, previously inaccessible data sources could entice further tourism businesses to make use of the VLCi platform, as this was one aspect that was criticized by one of the interviewees. Thus, in order to move beyond platform- or service-specific advancements, further systematic and automated integration and exchange of touristic data will be needed. Another desired development would be the integration of company databases with national organizations such as the INE. Destination-internal data exchange would also eliminate the need for organizations to rely on costly external data providers, which could be a hindrance. Instead, these analytics companies would deliver their insights to the platforms, which could ultimately put an end to the fragmentation of the data market. Last but not least, all of these efforts will only be worthwhile if the data platform is actually known and used by most tourism entities. Even though it could take some time, this endeavor has the potential to be fruitful in the long run as Big Data holds great promises.

CONCLUSION

Big Data is an emergent field that is currently transforming tourism philosophy and practice. The purpose of this research was to uncover the potentials and obstacles that arise within Smart Tourism Destinations, in this case Valencia, when leveraging Big Data. Seven semi-structured interviews with tourism stakeholders from the public and private sectors were conducted to study these factors. Several themes were discovered and their implications, as well as interrelations, were addressed during the analysis. As a result of the findings, one could say that Big Data holds benefits for both the supply and demand side of tourism as destination and tourism business managers are able to take advantage of the vast amount of data acquired to make overall better-informed decisions and improve the visitor experience by adding value through personalization. The more data an organization has, the better its situational awareness, enabling it to determine customer behaviour in terms of mobility, sentiment and purchasing patterns. In the long-term, Big Data will especially be helpful in uncertain times as it supports crisis response and management through the use of Big Data-enabled predictions to adapt to changes in the environment and demand. Hence, destinations and the related businesses may remain more adaptable and resilient as part of better strategies and decisions. Further, contemporary tourism issues such as overtourism and sustainability may be addressed by quantifying the economic, ecological and social impact of tourism through the collection of Big Data.

While Big Data promises a vast range of applications and information, the findings have shown that the reality is much more pragmatic as STDs and tourism enterprises face significant hurdles when it comes to implementing Big Data. Changes in organizational cultures, the need for financial investment and technologies, and having the proper skills appear to be some of the most significant obstacles, whilst simultaneously being requirements for utilizing Big Data. To be accepted, Big Data needs to be understood by everyone in the organization, at all levels, along with the added value it offers, why it is necessary, what data to use in which cases, and how it can be used to generate revenue and improve products and services. The transformation, however, does not stop at the internal level. New technologies, such as Big Data, necessitate a transformational process at all levels, including the corporate culture and mindset, and must be guided by a holistic and comprehensive strategic plan. In the end, solely having technology and putting it to use does not make a destination, industry, or business smart. It is not a matter of doing the same procedures with new technical applications but of revolutionizing tourism management. A destination will only be smart if each tourism stakeholder adopts these smart technologies and is embedded in the smart destination management. This transition will take time, strategy, and resources, as well as public and private capacity development, and is inextricably linked to the use of Big Data. Overall, numerous obstacles must be overcome before the potentials may be realized. At the same time, some of the aforementioned obstacles might also represent opportunities in terms of application possibilities for Big Data.

Some aspects that one should be aware of when employing Big Data in a STD context are to consider, for example, outsourcing data infrastructure and insights to specialized companies whilst internalizing basic data analysis skills. In order to derive value and insights from Big Data, even SMEs should have at least one employee specialised in data analysis in their team. Digitalization in general is an essential first step that needs to be taken before advocating advanced technological solutions such as Big Data. In the case of Spain, EU initiatives like the Digital Europe Program could help in accomplishing so. Collaboration is another element required to defeat the challenges of Big Data technologies. To achieve seamless information sharing among various entities inside STDs, data must be made available without charge as to remove obstacles. The interviewees recommended that a national open data dashboard and data gathering infrastructure should be established first, followed by data collection on the national, regional as well as local tourism area levels. Automated data collection could facilitate data sharing as well as the accessibility of data. This would also avoid duplication of efforts when tourism businesses or associations feel the need to additionally collect data in order to have local statistically more representative data or due to a lack of standards in data definitions or collection methods. The implementation of yet-to-be-established methodological standards would permit data sharing, integration, and cross-destination comparisons and guarantee the trustworthiness of results. In general, high-quality data, veracity, and confidentiality must be ensured. Abundant or contradictory data will otherwise waste resources and time and lead to unfavorable business decisions. Further, government statistics agencies should get more involved in the use and dissemination of new data sources.

Valencia as an STD has already recognized the need for technological solutions in helping to understand correlations between various variables and has taken the first steps on a destination level, as evidenced by the creation of value-added experiences through the AlterEco app and the collection, integration, and analysis of different types of data, such as done with the Valencia MOBILIS card. It is noteworthy that despite the relatively simple and small steps that Valencia has taken, the technological advancements of this particular STD still remain to be a best practice for exploiting Big Data in current times. Yet, the system ought to be improved by the integration of existing data platforms, next to business data, and the promotion of the SIT's capabilities in order to encourage the widespread use of the platform. The adaptation of a tourism intelligence system similar to the one in Valencia would be desirable for other destinations striving toward becoming STDs through the utilization of Big Data. At the same time, this may reduce the reluctance and entice further tourism businesses to start making use of Big Data in their operations. Also projects like HOSBEC's Biontrend may become the holy grail in terms of collaborative tools, where entities provide data that are combined to establish a foundation for efficient decision-making based on data rather than expertise or intuition.

From a theoretical standpoint, the report has contributed to the existing understanding of the notions of Big Data and Smart Tourism Destinations by defining and conceptualizing the

terms, which could aid in expanding the understanding of tourism businesses. Owing to the lack of research on the possible capabilities and impacts of Big Data technologies, another important contribution of this research is a review of Big Data applications in the STD context. This study advanced the discussion by examining and outlining the primary benefits and obstacles that STDs must overcome in order to successfully utilize Big Data and pinpointed considerations for the further implementation of this technology. Next to opportunities and challenges, the analysis also brought forward crucial enablers and identified a new challenge, namely organizational change, willingness, and understanding of the value of Big Data, which adds to the conceptualisation. This research therefore contributed to the limited tourism literature on the linkage of the topics of data analytics and Smart Tourism Destinations and looked at it from a more critical stance as it did not solely evaluate its benefits. All in all, this study helped to advance the theoretical framework surrounding Big Data utilization in STDs and to validate previous studies.

Thus, the findings at this stage of research can be utilized to assist local, regional and national undertakings aimed at Smart Tourism Destinations' strategies involving the use of Big Data. The previous findings may be helpful for DMOs and business managers as they provide insights on factors to consider when determining whether and how to incorporate Big Data. The identified key aspects for the implementation of Big Data could thus act as a guideline for other destinations interested in following a similar path. Knowing the dangers associated with Big Data adoption, tourism actors may make a more informed decision about how to engage in this technology, ensuring that the interests of all stakeholders are taken into consideration. Nonetheless, one must be aware that there is not one right way for STDs and the respective tourism business to implement Big Data technologies as the context, infrastructure and overall system differs in each destination and entity.

While this research has explored the potential implications of Big Data for Smart Tourism Destinations and thereby addressed a knowledge gap, it does bear some limitations. One of the major drawbacks of this analysis is related to the methodology and the challenge of generalization. Only eight tourism professionals were interviewed in total, with most tourism subsectors (e.g., lodging or DMO) being represented by only one interviewee. In the current study, this can be explained due to time constraints and the availability of interview partners. An increase in the number of respondents would aid data validity. To provide quantitative generalizable results, focus groups or surveys could be used, which again come with their own limitations. Furthermore, because this study is based on a single case study that is located in an urban Western context, the results cannot be generalized or directly applied to other contexts. This is because structures may differ significantly and additional challenges, such as basic internet connectivity, support systems, and funding, may arise. As a result, the study's conclusions are more indicative than conclusive. Additionally, only individuals who were comfortable with being interviewed in English were chosen for the interviews. This could have been a hurdle for other Spanish businesses or institutions who wanted to share their experiences but were not comfortable with communicating in English. Furthermore,

depending on their degree of fluency, respondents may not have been able to express themselves as clearly in English as they could have in their mother language.

This study focused mainly on the implications of Big Data on the tourism supply-side and neglected the perspectives of visitors and residents. This limitation is a consequence of the study's scope and focus. Lastly, the researcher comes from a tourism background and is not a specialist in data science, which might have influenced the interpretation and understanding of the results. Collaboration with a data scientist would be advisable if the study were to be repeated. The findings also suggest that there is a need for the analysis to be replicated in different STDs, with special attention paid to regions with varying levels of innovativeness and technological readiness in order to gain a better understanding of Big Data's potentialities. It remains to be seen how the results alter when the same study is conducted in a destination that is not yet considered "smart" in Spain. Are there any differences that can be attributed to smartness? This is just one question that arose during the investigation. This study's exploratory approach thus provides a starting point for future research on the implications of Big Data in an STD context.

REFERENCE LIST

1. Achaerandio, R., Gallotti, G., Curto, J., Bigliani, R., & Maldonado, F. (2011). Análisis de las Ciudades Inteligentes en España. *IDC*, 1–46.

2. Adams, W. C. (2015). Conducting Semi-Structured Interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation* (pp. 492–505). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119171386.ch19

3. AENOR. (n.d.). *AENOR Certification Smart tourism destination 178501-178502*. Retrieved April 28, 2022, from https://www.en.aenor.com/certificacion/administracion-publica/destino-turistico-inteligente

4. Ahas, R., Aasa, A., Roose, A., Mark, Ü., & Silm, S. (2008). Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management*, *29*(3), 469–486. https://doi.org/10.1016/j.tourman.2007.05.014

5. Aiden, E., & Michel, J.-B. (2013, December 25). The Predictive Power of Big Data. *Newsweek*. https://www.newsweek.com/predictive-power-big-data-225125

6. Ajuntament de València. (n.d.). *Geoportal València*. Retrieved April 28, 2022, from https://geoportal.valencia.es/

7. Ajuntament de València, & Oficina d'Estadística. (2018). *Àrea Metropolitana de València*.

8. Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*, *6*(25), 15. https://doi.org/10.1186/s13174-015-0041-5

9. Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*, *58*(2), 175–191. https://doi.org/10.1177/0047287517747753

10. Alastal, A., & Badawy, U. (2020). Big Data in Smart Cities: Exploring Possibilities in Terms of Opportunities and Challenges. *Journal of Scholarly Publishing*.

11. Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology*, *22*(1), 3–21. https://doi.org/10.1080/10630732.2014.942092

12. Alvesson, M., & Sköldberg, K. (2000). *Reflexive Methodology: New Vistas for Qualitative Research.* SAGE Publications.

13. Ambrosini, V., & Bowman, C. (2009). What are dynamic capabilities and are they a useful construct in strategic management? *International Journal of Management Reviews*, *11*(1), 29–49. https://doi.org/10.1111/j.1468-2370.2008.00251.x

14. Andrejevic, M., & Burdon, M. (2014). Defining the Sensor Society. *Television & New Media*, *16*(1), 19–36. https://doi.org/10.1177/1527476414541552

15. Angelidou, M. (2014). Smart city policies: A spatial approach. *Cities*, *41*, 3–11. https://doi.org/10.1016/j.cities.2014.06.007

16. Ann Keller, S., Koonin, S. E., & Shipp, S. (2012). Big data and city living – what can it do for us? *Significance*, *9*(4), 4–7. https://doi.org/10.1111/j.1740-9713.2012.00583.x

17. Anuar, F. I., & Gretzel, U. (2011). *Privacy Concerns in the Context of Location-Based Services for Tourism*. 6. http://agrilife.org/ertr/files/2013/02/13.pdf

18. Ardagna, C. A., Ceravolo, P., & Damiani, E. (2016). Big data analytics as-a-service: Issues and challenges. *2016 IEEE International Conference on Big Data (Big Data)*, 3638–3644. https://doi.org/10.1109/BigData.2016.7841029

19. Ardito, L., Cerchione, R., Del Vecchio, P., & Raguseo, E. (2019). Big data in smart tourism: Challenges, issues and opportunities. *Current Issues in Tourism*, *22*(15), 1805–1809. https://doi.org/10.1080/13683500.2019.1612860

20. Arellano, A., Cámara, N., Carta, G., & García, J. R. (2021). *Spain | Digital skills: What are they and who has them?* BBVA Research. https://www.bbvaresearch.com/en/publicaciones/spain-digital-skills-what-are-they-and-who-has-them/

21. Arksey, H. (1996). Collecting data through joint interviews. *Social Research Update*, *15*. https://sru.soc.surrey.ac.uk/SRU15.html

22. Avelar, S. (2020). From a Smart City to a Smart Destination: A Case Study. In *Strategic Innovative Marketing and Tourism* (pp. 7–14). Springer. https://ideas.repec.org/h/spr/prbchp/978-3-030-36126-6_2.html

23. Babar, M., & Arif, F. (2017). Smart urban planning using Big Data analytics to contend with the interoperability in Internet of Things. *Future Generation Computer Systems*, *77*, 65–76. https://doi.org/10.1016/j.future.2017.07.029

24. Bagdonien, L., & Zemblyté, J. (2005). Service research: Advantages and limitations of quantitative and qualitative approaches. *Socialiniai Mokslai*, *4*(50), 26–37.

25. Baggio, R. (2016). Improving tourism statistics: Merging official records with big data. In M. Fuchs, W. Höpken, & M. Lexhagen (Eds.), *Big Data and Business Intelligence in the Travel and Tourism Industry* (pp. 89–92).

26. Baggio, R., & Cooper, C. (2010). Knowledge transfer in a tourism destination: The effects of a network structure. *The Service Industries Journal*, *30*(10), 1757–1771. https://doi.org/10.1080/02642060903580649

27. Baggio, R., Micera, R., & Del Chiappa, G. (2020). Smart tourism destinations: A critical reflection. *Journal of Hospitality and Tourism Technology*, *11*(3), 407–423. https://doi.org/10.1108/JHTT-01-2019-0011

28. Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, *68*, 101–115. https://doi.org/10.1016/j.tourman.2018.02.020

29. Baucells, N., Moreno, C., & Arce, R. M. (2016). Mapping Smart Cities Situation +CITIES: The Spanish Case. In E. Alba, F. Chicano, & G. Luque (Eds.), *Smart Cities* (pp. 95–105). Springer International Publishing. https://doi.org/10.1007/978-3-319-39595-1_10

30. Beath, C., Becerra-Fernandez, I., Ross, J., & Short, J. (2012). Finding Value in the Information Explosion. *MIT Sloan Management Review*, *53*(4), 18–20.

31. Beeton, S. (2005). The case study in tourism research: A multi-method case study approach. In B. W. Ritchie, P. Burns, & C. Palmer (Eds.), *Tourism research methods: Integrating theory with practice* (pp. 37–48). CABI Publishing. https://www.cabdirect.org/cabdirect/abstract/20053116157

32. Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, *11*(3), 369–386. https://doi.org/10.2307/248684

33. Bertocchi, D., Camatti, N., Giove, S., & van der Borg, J. (2020). Venice and Overtourism: Simulating Sustainable Development Scenarios through a Tourism Carrying Capacity Model. *Sustainability*, *12*(2), 512. https://doi.org/10.3390/su12020512

34. Bertot, J. C., & Choi, H. (2013). Big data and e-government: Issues, policies, and recommendations. *Proceedings of the 14th Annual International Conference on Digital Government Research*, 1–10. https://doi.org/10.1145/2479724.2479730

35. Bibri, S. E., & Krogstie, J. (2017). The core enabling technologies of big data analytics and context-aware computing for smart sustainable cities: A review and synthesis. *Journal of Big Data*, *4*(1), 38. https://doi.org/10.1186/s40537-017-0091-6

36. Bicing. (n.d.). *Home*. Bicing. Retrieved April 18, 2022, from https://www.bicing.barcelona/es

37. Boes, K., Borde, L., & Egger, R. (2015). The Acceptance of NFC Smart Posters in Tourism. In I. Tussyadiah & A. Inversini (Eds.), *Information and Communication Technologies in Tourism 2015* (pp. 435–447). Springer International Publishing.

38. Boes, K., Buhalis, D., & Inversini, A. (2015). Conceptualising Smart Tourism Destination Dimensions. In I. Tussyadiah & A. Inversini (Eds.), *Information and Communication Technologies in Tourism 2015* (pp. 391–403). Springer International Publishing. https://doi.org/10.1007/978-3-319-14343-9

39. Boes, K., Buhalis, D., & Inversini, A. (2016). Smart tourism destinations: Ecosystems for tourism destination competitiveness. *International Journal of Tourism Cities*, *2*(2), 108–124. https://doi.org/10.1108/IJTC-12-2015-0032

40. Bose, R. (2009). Advanced analytics: Opportunities and challenges. *Industrial Management & Data Systems*, *109*(2), 155–172. https://doi.org/10.1108/02635570910930073

41. Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

42. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

43. Braunhofer, M., & Ricci, F. (2017). Selective contextual information acquisition in travel recommender systems. *Information Technology & Tourism*, *17*(1), 5–29. https://doi.org/10.1007/s40558-017-0075-6

44. Brynjolfsson, E. (1993). The Productivity Paradox of Information Technology. *Communications of the ACM*, *36*(12), 66–77. https://doi.org/10.1145/163298.163309

45. Bryson, S., Kenwright, D., Cox, M., Ellsworth, D., & Haimes, R. (1999). Visually exploring gigabyte data sets in real time. *Communications of the ACM*, *42*(8), 82–90. https://doi.org/10.1145/310930.310977

46. Buckley, J. B. (2022, April 22). Venice to charge daytrippers up to €10 to enter in 2023. *CNN*. https://www.cnn.com/travel/article/venice-entry-reservation-fee/index.html

47. Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management*, *21*(1), 97–116. https://doi.org/10.1016/S0261-5177(99)00095-3

48. Buhalis, D. (2015, February 10). Working Definitions of Smartness and Smart Tourism Destination. *Working Definitions of Smartness and Smart Tourism Destination*. http://buhalis.blogspot.com/2014/12/working-definitions-of-smartness-and.html

49. Buhalis, D., & Amaranggana, A. (2013). Smart Tourism Destinations. In Z. Xiang & I. Tussyadiah (Eds.), *Information and Communication Technologies in Tourism 2014* (pp. 553–564). Springer International Publishing. https://doi.org/10.1007/978-3-319-03973-2_40

50. Buhalis, D., & Amaranggana, A. (2015). Smart Tourism Destinations Enhancing Tourism Experience Through Personalisation of Services. In I. Tussyadiah & A. Inversini (Eds.), *Information and Communication Technologies in Tourism 2015* (pp. 377–389). Springer International Publishing. https://doi.org/10.1007/978-3-319-14343-9_28

51. Buhalis, D., & Foerste, M. (2015). SoCoMo Marketing for Travel and Tourism: Empowering Co-Creation of Value. *Journal of Destination Marketing & Management*, *4*(3), 151–161. https://doi.org/10.1016/j.jdmm.2015.04.001

52. Buhalis, D., & Law, R. (2008). 20 Years on and 10 Years After the Internet—The State of eTourism Research. *Tourism Management*, *29*(4), 609–623. https://doi.org/10.1016/j.tourman.2008.01.005

53. Burkart, A. J., & Medlik, S. (1981). *Tourism: Past, Present and Future*. Elsevier Science & Technology Books.

54. Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82. https://doi.org/10.1080/10630732.2011.601117

55. Celdrán Bernabeu, M. A., Mazón, J.-N., Giner Sanchez, D., & Ivars Baidal, J. A. (2016). *Big Data and Smart Tourism Destinations: Challenges and opportunities from an industry perspective*. 1–17.

56. Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, *63*, 67–80. https://doi.org/10.1016/j.dss.2013.08.008

57. Charmaz, K. (2006). *Constructing grounded theory*. Sage Publications.

58. Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

59. Chen, Y., Ardila-Gomez, A., & Frame, G. (2017). Achieving Energy Savings by Intelligent Transportation Systems Investments in the Context of Smart Cities. *Transportation Research Part D: Transport and Environment*, *54*, 381–386. https://doi.org/10.1016/j.trd.2017.06.008

60. Choe, Y., & Fesenmaier, D. R. (2017). The Quantified Traveller: Implications for Designing Tourism Systems. In *Analytics in Smart Tourism Design* (pp. 65–77). Springer.

61. Cimbaljević, M., Stankov, U., & Pavluković, V. (2019). Going beyond the traditional destination competitiveness – reflections on a smart destination in the current research. *Current Issues in Tourism*, *22*(20), 2472–2477. https://doi.org/10.1080/13683500.2018.1529149

62. CIS. (n.d.). *Barómetro de la Comunidad Valenciana* [Map]. Retrieved April 28, 2022, from https://www.cis.es/cis/export/sites/default/-Archivos/Marginales/1840_1859/1857/Es1857mar_PDF_enlace.html

63. Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 3rd ed* (pp. xv, 379). Sage Publications, Inc. https://doi.org/10.4135/9781452230153

64. Creswell, J. (2003). *Research Design: Qualitative, Quantitative, and Mixed-Method Approaches* (2nd ed.). SAGE Publications, Inc.

65. Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications.

66. Crouch, G. I., & Ritchie, J. R. B. (1999). Tourism, Competitiveness, and Societal Prosperity. *Journal of Business Research*, *44*(3), 137–152. https://doi.org/10.1016/S0148-2963(97)00196-3

67. d'Amore, M., Baggio, R., & Valdani, E. (2015). A Practical Approach to Big Data in Tourism: A Low Cost Raspberry Pi Cluster. In I. Tussyadiah & A. Inversini (Eds.), *Information and Communication Technologies in Tourism 2015* (pp. 169–181). Springer. https://doi.org/10.1007/978-3-319-14343-9_13

68. Davenport, T. H. (2019). *At the Big Data Crossroads: Turning towards a smarter travel experience* [Data set]. Amadeus. https://doi.org/10.1287/8943f842-86f8-4d42-9a64-9a7cd07b31f5

69. Davenport, T. H., & Patil, D. J. (2012, October 1). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*. https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

70. Day, G. S. (2011). Closing the Marketing Capabilities Gap. *Journal of Marketing*, *75*, 183–195.

71. De Filippi, P. (2015). Community Mesh Networks: Citizens Participation in the Deployment of Smart Cities. In A. Vesco & F. Ferrero (Eds.), *Social, Economic, and Environmental Sustainability in the Development of Smart Cities* (pp. 298–314). IGI Global. https://doi.org/10.4018/978-1-4666-8282-5.ch014

72. DeCarlo, M. (2018). Sampling in qualitative research. In *Scientific Inquiry in Social Work* (pp. 271–278). Open Social Work Education. https://open.umn.edu/opentextbooks/textbooks/591

73. Dekay, S. H. (2012). How large companies react to negative Facebook comments. *Corporate Communications: An International Journal*, *17*(3), 289–299. https://doi.org/10.1108/13563281211253539

74. Del Chiappa, G., & Baggio, R. (2015). Knowledge transfer in smart tourism destinations: Analyzing the effects of a network structure. *Journal of Destination Marketing & Management*, *4*(3), 145–150. https://doi.org/10.1016/j.jdmm.2015.02.001

75. Del Chiappa, G., & Dall'Aglio, S. (2012). Factors influencing Travellers' e-Ratings and e-Complaints about Hotel Services: Insights from an Italian Tourism Destination. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and Communication Technologies in Tourism 2012* (pp. 448–459). Springer. https://doi.org/10.1007/978-3-7091-1142-0_39

76. Del Vecchio, P., Mele, G., Ndou, V., & Secundo, G. (2018). Creating value from Social Big Data: Implications for Smart Tourism Destinations. *Information Processing & Management*, *54*(5), 847–860. https://doi.org/10.1016/j.ipm.2017.10.006

77. Demunter, C., European Commission, & Statistical Office of the European Union. (2017). *Tourism statistics: Early adopters of big data? : 2017 edition.* Publications Office of the European Union. http://dx.publications.europa.eu/10.2785/762729

78. Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The Discipline and Practice of Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 1–32). Sage Publications Ltd.

79. Desjardins, J. (2019, April 17). *How much data is generated each day?* World Economic Forum. https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/

80. Dexeus, C. R. (2021). *Smart Tourism Destinations: Boosting digitalization and tourism innovation*. SEGITTUR. https://www.intelligentcitieschallenge.eu/sites/default/files/2021-07/ICC_2%20ICC%20City%20Lab_Day%202_Smart%20Tourism%20Destinations_Carlos%20Romero.pdf

81. Directorate-General for Communication. (2021, December 6). *Flash Eurobarometer 499: Attitudes of Europeans towards tourism—Data Europa EU*. https://data.europa.eu/data/datasets/s2283_499_eng?locale=en

82. Dodds, R., & Butler, R. W. (Eds.). (2020). *Overtourism: Issues, realities and solutions* (Vol. 15). Walter de Gruyter. https://doi.org/10.1080/1743873X.2019.1701818

83. Domo. (n.d.). *Data Never Sleeps 9.0*. Retrieved April 13, 2022, from https://www.domo.com/learn/infographic/data-never-sleeps-9

84. Doran, J., Ryan, G., Bourke, J., & Crowley, F. (2020). In-house or outsourcing skills: How best to manage for innovation? *International Journal of Innovation Management*, *24*(01), 2050010. https://doi.org/10.1142/S1363919620500103

85. Dredge, D., Phi, G., Mahadevan, R., Meehan, E., & Popescu, E. S. (2018). *Digitalisation in Tourism: In-depth analysis of challenges and opportunities*. Executive Agency for Small and Medium-sized Enterprises (EASME) Virtual Tourism Observatory. Aalborg University, Copenhagen. https://ehrl.ee/wp-content/uploads/2019/01/Digitalisation-in-Tourism-Report-2018.pdf

86. Easterby-Smith, M., Thorpe, R., & Jackson, P. R. (2015). *Management & Business Research* (5th ed.). Sage Edge.

87. Eberhart, A. (2015). Thinking green, buying green? Drivers of pro - Environmental purchasing behavior. *Journal of Consumer Marketing*, *32*, 167–175. https://doi.org/10.1108/JCM-10-2014-1179

88. Egger, R. (2013). The impact of near field communication on tourism. *Journal of Hospitality and Tourism Technology*, *4*(2), 119–133. https://doi.org/10.1108/JHTT-04-2012-0014

89. Ejimabo, N. O. (2015). The effective research process: Unlocking the advantages of ethnographic strategies in the qualitative research methods. *European Scientific*

Journal, *ESJ*, *11*(23), Article 23. https://eujournal.org/index.php/esj/article/view/6090

90. Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazi, A., Bowman, T., Suri, V., Tsou, A., Weingart, S., & Sugimoto, C. (2015). Big Data, Bigger Dilemmas: A Critical Review. *Journal of the Association for Information Science and Technology*, *66*(8), 1523–1545. https://doi.org/10.1002/asi.23294

91. El Alaoui, I., Gahi, Y., Messoussi, R., Todoskoff, A., & Kobi, A. (2017). *Big Data Analytics: A Comparison of Tools and Applications*. 587. https://doi.org/10.1007/978-3-319-74500-8_54

92. Erl, T., Khattak, W., & Buhler, P. (2016). *Big data fundamentals: Concepts, drivers & techniques*. Prentice Hall.

93. European Commission. (n.d.-a). *European Capitals of Smart Tourism*. Retrieved April 7, 2022, from https://smart-tourism-capital.ec.europa.eu/index_de

94. European Commission. (n.d.-b). *Smart cities*. Retrieved April 4, 2022, from https://ec.europa.eu/info/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en

95. European Commission. (n.d.-c). *The Digital Economy and Society Index (DESI) | Shaping Europe's digital future*. Retrieved May 23, 2022, from https://digital-strategy.ec.europa.eu/en/policies/desi

96. European Commission. (2020a). *Leading examples of Smart Tourism Practices in Europe*. https://smart-tourism-capital.ec.europa.eu/cities/leading-examples-smart-tourism-practices-europe_en

97. European Commission. (2020b, March 5). *Commission's agreement with collaborative economy platforms*. https://ec.europa.eu/commission/presscorner/detail/en/ip_20_194

98. European Commission. (2021, January 21). *Building a data economy—Brochure*. https://digital-strategy.ec.europa.eu/en/library/building-data-economy-brochure

99. European Commission, Directorate-General for Communication, C. and T. (European C., Cattaneo, G., Micheletti, G., Glennon, M., La Croce, C., & Mitta, C. (2020). *The European data market monitoring tool: Key facts & figures, first policy conclusions, data landscape and quantified stories : d2.9 final study report Executive summary*. Publications Office of the European Union. https://data.europa.eu/doi/10.2759/837980

100. European Commission Directorate-General for Communications Networks, Content and Technology. (n.d.-a). *A European Strategy for data*. Retrieved April 18, 2022, from https://digital-strategy.ec.europa.eu/en/policies/strategy-data

101. European Commission Directorate-General for Communications Networks, Content and Technology. (n.d.-b). *Europe's Digital Decade: Digital targets for 2030* [Text]. Retrieved April 18, 2022, from https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en

102. A New Industrial Strategy for Europe, COM(2020) 102 final (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593086905382&uri=CELEX:52020DC0102

103. European Commission Directorate-General for Communications Networks, Content and Technology & Industry Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs. (2022). *Transition pathway for tourism*. Publications Office of the European Union. https://data.europa.eu/doi/10.2873/344425

104. Exceltur. (2018). *Comunitat Valenciana—IMPACTUR 2017*. https://www.exceltur.org/wp-content/uploads/2018/12/IMPACTUR-Comunidad-Valenciana-2017.pdf

105. Ezzy, D. (2013). *Qualitative Analysis* (1st ed.). Routledge. https://doi.org/10.4324/9781315015484

106. Falconer, G., & Mitchell, S. (2012). *A Systematic Process for Enabling Smart+Connected Communities*. Cisco Internet Business Solutions Group.

107. Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data analysis. *National Science Review*, *1*(2), 293–314. https://doi.org/10.1093/nsr/nwt032

108. Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, *57*(6), 74–81. https://doi.org/10.1145/2602574

109. Figueredo, M., Ribeiro, J., Cacho, N., Thome, A., Cacho, A., Lopes, F., & Araujo, V. (2018). From Photos to Travel Itinerary: A Tourism Recommender System for Smart Tourism Destination. *2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService)*, 85–92. https://doi.org/10.1109/BigDataService.2018.00021

110. Formia, S., Vivas, H. L., Cambarieri, M. G., Muñoz Abbate, H., García Martínez, N., & Petroff, M. (2015). Tratamiento de Grandes Volúmenes de Datos en Ciudades Inteligentes Una Propuesta de Big Data con NoSQL. *Universidad Nacional de Río Negro*. http://rid.unrn.edu.ar/handle/20.500.12049/150

111. ForwardKeys. (n.d.-a). *About the company*. ForwardKeys. Retrieved May 27, 2022, from https://forwardkeys.com/forwardkeys-about-the-company/

112. ForwardKeys. (n.d.-b). *Democratising data for travel operators & partners—Visit Valencia*. Retrieved May 27, 2022, from https://forwardkeys.com/case/visit-valencia-case-study/?portfolioCats=47

113. Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics, 165*, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031

114. Framke, W. (2002). The Destination as a Concept: A Discussion of the Business-related Perspective versus the Socio-cultural Approach in Tourism Theory. *Scandinavian Journal of Hospitality and Tourism*, *2*(2), 92–108. https://doi.org/10.1080/15022250216287

115. Frederiksen, L. (2012). Big Data. *Public Services Quarterly*, *8*(4), 345–349. https://doi.org/10.1080/15228959.2012.730411

116. Fuchs, M., & Höpken, W. (2011). E-Business horizons in the tourism industry - Challenges for research and practice. In K. Sidali, A. Spiller, & B. Schulze (Eds.), *Food, agriculture and tourisms – Linking local gastronomy and rural tourism* (pp. 140–160). Springer. https://link.springer.com/chapter/10.1007/978-3-642-11361-1_10

117. Fuchs, M., Höpken, W., & Lexhagen, M. (2014). Big data analytics for knowledge generation in tourism destinations – A case from Sweden. *Journal of Destination Marketing & Management*, *3*(4), 198–209. https://doi.org/10.1016/j.jdmm.2014.08.002

118. Fundación Valenciaport. (n.d.). *About Us*. Fundación Valenciaport. Retrieved May 27, 2022, from https://www.fundacion.valenciaport.com/en/about-us/

119. Gajdošík, T. (2019). Big Data Analytics in Smart Tourism Destinations. A New Tool for Destination Management Organizations? In V. Katsoni & M. Segarra-Oña (Eds.), *Smart Tourism as a Driver for Culture and Sustainability* (pp. 15–33). Springer International Publishing. https://doi.org/10.1007/978-3-030-03910-3_2

120. Gallego, I., & Font, X. (2021). Changes in air passenger demand as a result of the COVID-19 crisis: Using big data to inform tourism policy. *Journal of Sustainable Tourism*, *29*(9), 1470–1489.

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

122. Generalitat Valenciana. (n.d.). *Turisme Comunitat Valenciana*. ¿Quiénes Somos?: Organización y Organigrama. Retrieved May 27, 2022, from https://www.turisme.gva.es/opencms/opencms/turisme/es/contents/conselleria/organigrama/organigrama.html?tam=&menu_id=1

123. George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big Data and Data Science Methods for Management Research. *Academy of Management Journal*, *59*(5), 1493–1507. https://doi.org/10.5465/amj.2016.4005

124. Georgulas, N. (1970). Tourist destination features. *Journal of Town Planning Institute*, *56*, 442–446.

125. Ghose, A., Ipeirotis, P. G., & Li, B. (2012). Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. *Marketing Science*, *31*(3), 493–520. https://doi.org/10.1287/mksc.1110.0700

126. Gibson, G., Timlin, A., Curran, S., & Wattis, J. (2004). The scope for qualitative methods in research and clinical trials in dementia. *Age and Ageing*, *33*(4), 422–426. https://doi.org/10.1093/ageing/afh136

127. Gibson, L. (2010). *Using Email Interviews*. ERSC National Center for Research Methods. https://eprints.ncrm.ac.uk/id/eprint/1303/1/09-toolkit-email-interviews.pdf

128. Gibson, W., & Brown, A. (2009). *Working with Qualitative Data*. SAGE Publications, Ltd. https://doi.org/10.4135/9780857029041

129. Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Milanović, N., & Meijers, E. (2007). *Smart cities—Ranking of European medium-sized cities*. Vienna University of Technology.

130. Given, L. M., & Saumure, K. (2008). Nonprobability sampling. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (1st ed., pp. 562–563). Sage Publications.

131. Gobble, M. M. (2013). Big Data: The Next Big Thing in Innovation. *Research-Technology Management*, *56*(1), 64–67. https://doi.org/10.5437/08956308X5601005

132. Google Trends. (n.d.). *Google Trends*. Google Trends. Retrieved March 26, 2022, from https://trends.google.de/trends/explore?date=today%205-y&q=smart%20tourism

133. Google Trends. (2022). *Big Data*. Google Trends. https://trends.google.de/trends/explore?q=big%20data&geo=DE

134. Gorini, M. (n.d.). *How Big Data Technologies Can Improve Tourism*. Bismart. Retrieved April 28, 2022, from http://blog.bismart.com/en/big-data-technologies-tourism

135. Gretzel, U. (2011). Intelligent systems in tourism: A Social Science Perspective. *Annals of Tourism Research*, *38*(3), 757–779. https://doi.org/10.1016/j.annals.2011.04.014

136. Gretzel, U. (2018). From smart destinations to smart tourism regions. *Investigaciones Regionales – Journal of Regional Research*, *42*, 171–184.

137. Gretzel, U., & Koo, C. (2021). Smart tourism cities: A duality of place where technology supports the convergence of touristic and residential experiences. *Asia Pacific Journal of Tourism Research*, *26*(4), 352–364. https://doi.org/10.1080/10941665.2021.1897636

138. Gretzel, U., Reino, S., Kopera, S., & Koo, C. (2015). Smart Tourism Challenges. *Journal of Tourism*, *16*(1), 41–47.

139. Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: Foundations and developments. *Electronic Markets*, *25*(3), 179–188. https://doi.org/10.1007/s12525-015-0196-8

140. Gretzel, U., Werthner, H., Koo, C., & Lamsfus, C. (2015). Conceptual foundations for understanding smart tourism ecosystems. *Computers in Human Behavior*, *50*, 558–563. https://doi.org/10.1016/j.chb.2015.03.043

141. Gretzel, U., Zhong, L., & Koo, C. (2016). Application of smart tourism to cities. *International Journal of Tourism Cities*, *2*(2). https://doi.org/10.1108/IJTC-04-2016-0007

142. Guba, E. G. (1990). *The paradigm dialog* (p. 424). Sage Publications, Inc.

143. Guba, E. G., & Lincoln, Y. S. (1994). Competing Paradigms in Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (3rd ed., pp. 105–117). SAGE Publications.

144. Guo, Y., Liu, H., & Chai, Y. (2014). The embedding convergence of smart cities and tourism internet of things in China: An advance perspective. *Advances in Hospitality and Tourism Research*, *2*(1), 54–69.

145. Halevi, G., & Moed, H. (2012). The evolution of big data as a research and scientific topic: Overview of the literature. *Research Trends*, *1*(30), 3–6.

146. Hall, R., Bowerman, B., Braverman, J., Taylor, J., Todosow, H., & Wimmersperg, U. (2000). The vision of a smart city. *Brookhaven National Laboratory*.

147. Hammer, M. (1990, July 1). Reengineering Work: Don't Automate, Obliterate. *Harvard Business Review*, *68*(4), 104–112.

148. Harding, J. (2015). *Identifying Themes and Coding Interview Data: Reflective Practice in Higher Education*. SAGE Publications, Ltd. https://doi.org/10.4135/9781473942189

149. Harrison, C., Eckman, B., Hamilton, R., Hartswick, P., Kalagnanam, J., Paraszczak, J., & Williams, P. (2010). Foundations for Smarter Cities. *IBM Journal of Research and Development*, *54*(4), 1–16. https://doi.org/10.1147/JRD.2010.2048257

150. Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, *47*, 98–115. https://doi.org/10.1016/j.is.2014.07.006

151. Hashem, I., Chang, V., Anuar, N., S., A., Yaqoob, I., Gani, A., Ahmed, E., & Chiroma, H. (2016). The Role of Big Data in Smart City. *International Journal of Information Management*, *36*(5). https://doi.org/10.1016/j.ijinfomgt.2016.05.002

152. Hautamäki, A. (2020). *Viewpoint Relativism: A New Approach to Epistemological Relativism based on the Concept of Points of View* (Vol. 419). Springer International Publishing. https://doi.org/10.1007/978-3-030-34595-2

153. Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, *41*(3), 260–271. https://doi.org/10.1080/15230406.2014.890072

154. Hawkins, J. (2018). The Practical Utility and Suitability of Email Interviews in Qualitative Research. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2018.3266

155. Hendrik, H., & Perdana, D. H. F. (2014). Trip Guidance: A Linked Data Based Mobile Tourists Guide. *Advanced Science Letters*, *20*(1), 75–79. https://doi.org/10.1166/asl.2014.5285

156. Hernández-Muñoz, J. M., Vercher, J. B., Muñoz, L., Galache, J. A., Presser, M., Gómez, L. A. H., & Pettersson, J. (2011). Smart Cities at the forefront of the future internet. In *The Future Internet: Future Internet Assembly 2011: Achievements and Technological Promise* (pp. 447–462). Springer Science+Business Media B.V. http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-20090

157. Hjalager, A.-M. (2002). Repairing innovation defectiveness in tourism. *Tourism Management*, *23*(5), 465–474. https://doi.org/10.1016/S0261-5177(02)00013-4

158. Hjalager, A.-M. (2010). A review of innovation research in tourism. *Tourism Management*, *31*(1), 1–12. https://doi.org/10.1016/j.tourman.2009.08.012

159. Hollands, R. (2014). Critical interventions into the corporate smart city. *Cambridge Journal of Regions, Economy and Society*, *8*(1), 61–77. https://doi.org/10.1093/cjres/rsu011

160. Hollands, R. G. (2008). Will the real smart city please stand up? *City*, *12*(3), 303–320. https://doi.org/10.1080/13604810802479126

161. HOSBEC. (n.d.). *Sobre Hosbec*. Retrieved May 27, 2022, from https://hosbec.com/hosbec/sobre-nosotros/

162. Howie, F. (2003). Managing the tourist destination. *Managing the Tourist Destination*. https://www.cabdirect.org/cabdirect/abstract/20033102971

163. Hu, H., Wen, Y., Chua, T.-S., & Li, X. (2014). Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. *IEEE Access*, *2*, 652–687. https://doi.org/10.1109/ACCESS.2014.2332453

164. Hu, Y.-H., Chen, Y.-L., & Chou, H.-L. (2017). Opinion mining from online hotel reviews – A text summarization approach. *Information Processing & Management*, *53*(2), 436–449. https://doi.org/10.1016/j.ipm.2016.12.002

165. Hwang, J., Park, H.-Y., & Hunter, W. C. (2015). Constructivism in Smart Tourism Research: Seoul Destination Image. *Asia Pacific Journal of Information Systems*, *25*(1), 163–178. https://doi.org/10.14329/apjis.2015.25.1.163

166. IBM. (2014). *Integrating and governing big data*. IBM Software.

167. IBM. (2015, October 1). *IBM builds a smarter planet*. http://www.ibm.com/smarterplanet/us/en/

168. IDOM. (n.d.). Impulso Valencia Smart City—VLCi. *IDOM*. Retrieved April 7, 2022, from https://www.idom.com/en/project/impulso-vlci-valencia-smart-city/

169. Instituto Nacional de Estadística. (n.d.). *INEbase*. https://www.ine.es/dynt3/inebase/es/index.htm?padre=2382&capsel=3149

170. Instituto Nacional de Estadística. (2022a). *Cifras PYME. Datos mayo 2022*. https://industria.gob.es/es-es/estadisticas/Cifras_PYME/CifrasPYME-mayo2022.pdf

171. Instituto Nacional de Estadística. (2022b). *Cuenta Satélite del Turismo de España. Revisión estadística 2019*. https://www.ine.es/prensa/cst_2020.pdf

172. Instituto Nacional de Estadística. (2022c). *Estadística de Movimientos Turísticos en Fronteras- FRONTUR*. https://www.ine.es/en/daco/daco42/frontur/frontur1221_en.pdf

173. INVATTUR. (n.d.). *Modelo DTI de la Comunitat Valenciana*. Retrieved April 7, 2022, from https://invattur.es/modelo-destinos-turisticos-inteligentes.html

174. INVATTUR, & Territorio Creativo. (2015). *Big Data: Retos y oportunidades para el turismo*. https://www.thinktur.org/media/Big-Data.-Retos-y-oportunidades-para-el-turismo.pdf

175. ISO. (n.d.). *ISO 37120:2018*. ISO 37120:2018 Sustainable Cities and Communities — Indicators for City Services and Quality of Life. Retrieved April 7, 2022, from https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/84/68498.html

176. ITU. (n.d.). *United for Smart Sustainable Cities (U4SSC) – United for Smart Sustainable Cities*. Retrieved April 28, 2022, from https://u4ssc.itu.int/

177. Iunius, R., Cismaru, L., & Foris, D. (2015). Raising Competitiveness for Tourist Destinations through Information Technologies within the Newest Tourism Action Framework Proposed by the European Commission. *Sustainability*, *7*(9), 12891–12909. https://doi.org/10.3390/su70912891

178. Ivars Baidal, J. A. (2014). *Destinos Turísticos Inteligentes Comunitat Valenciana*. INVATTUR. https://invattur.es/uploads/entorno_37/ficheros/626905489baf21583379105.pdf

179. Ivars-Baidal, J. A., Celdrán-Bernabeu, M. A., Mazón, J.-N., & Perles-Ivars, Á. F. (2019). Smart destinations and the evolution of ICTs: A new scenario for destination management? *Current Issues in Tourism*, *22*(13), 1581–1600. https://doi.org/10.1080/13683500.2017.1388771

180. IVE. (n.d.). *ALTER ECO: Alternative tourist strategies to enhance the local sustainable development of tourism by promoting Mediterranean Identity*. Retrieved March 31, 2022, from https://www.five.es/project/alter-eco/

181. Jacobsen, J. Kr. S., & Munar, A. M. (2012). Tourist information search and destination choice in a digital age. *Tourism Management Perspectives*, *1*, 39–47. https://doi.org/10.1016/j.tmp.2011.12.005

182. Jani, D., Jang, J.-H., & Hwang, Y.-H. (2014). Big Five Factors of Personality and Tourists' Internet Search Behavior. *Asia Pacific Journal of Tourism Research*, *19*(5), 600–615. https://doi.org/10.1080/10941665.2013.773922

183. Jasrotia, A., & Gangotia, A. (2018). Smart cities to smart tourism destinations: A review paper. *Journal of Tourism Intelligence and Smartness*, *1*, 47–56.

184. Jennings, G. R. (2015). *Qualitative Research for University Sector: Paradigms that inform Research*. Imagine Consulting Group International Pty Ltd.

185. Ji, C., Li, Y., Qiu, W., Awada, U., & Li, K. (2012). Big Data Processing in Cloud Computing Environments. *2012 12th International Symposium on Pervasive Systems, Algorithms and Networks*, 17–23. https://doi.org/10.1109/I-SPAN.2012.9

186. Jiménez, E. (2015). Las Ciudades Inteligentes y los Big Data. *Universidad Central de Venezuela*, 11.

187. Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and Challenges of Big Data Research. *Big Data Research*, *2*(2), 59–64. https://doi.org/10.1016/j.bdr.2015.01.006

188. Johannessen, A., Christoffersen, L., & Tufte, P. A. (2020). *Forskningsmetode for økonomisk-administrative fag*. Abstrakt Forlag.

189. Jones, M. (2012, June 27). *Big Data Is a 'New Natural Resource,' IBM Says*. GovTech. https://www.govtech.com/archive/Big-Data-Is-a-New-Natural-Resource-IBM-Says.html

190. Jovicic, D. Z. (2019). From the traditional understanding of tourism destination to the smart tourism destination. *Current Issues in Tourism*, *22*(3), 276–282. https://doi.org/10.1080/13683500.2017.1313203

191. Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*(7), 2561–2573. https://doi.org/10.1016/j.jpdc.2014.01.003

192. Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: Issues, challenges, tools and Good practices. *2013 Sixth International Conference on Contemporary Computing (IC3)*, 404–409. https://doi.org/10.1109/IC3.2013.6612229

193. Khalid, U., Okafor, L. E., & Burzynska, K. (2021). Does the size of the tourism sector influence the economic policy response to the COVID-19 pandemic? *Current Issues in Tourism*, *24*(19), 2801–2820. https://doi.org/10.1080/13683500.2021.1874311

194. Khan, M. A., Uddin, M. F., & Gupta, N. (2014). *Seven V's of Big Data—Understanding Big Data to extract Value*. 5.

195. Khan, Z., Anjum, A., & Kiani, S. L. (2013). Cloud Based Big Data Analytics for Smart Future Cities. *2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*, 381–386. https://doi.org/10.1109/UCC.2013.77

196. Khomsi, M. R., & Bedard, F. B. (2016). From smart city to smart destination. The case of three Canadian cities. *Journal of Tourism Research*, *6*(2), 69–74.

197. Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-data applications in the government sector. *Communications of the ACM*, *57*(3), 78–85. https://doi.org/10.1145/2500873

198. Kiron, D. (2013). Organizational Alignment is Key to Big Data Success. *MIT Sloan Management Review*, *54*(3). https://sloanreview.mit.edu/article/organizational-alignment-is-key-to-big-data-success/

199. Kitchin, R., & Lauriault, T. P. (2015). Small data in the era of big data. *GeoJournal*, *80*(4), 463–475. https://doi.org/10.1007/s10708-014-9601-7

200. Komninos, N., Pallot, M., & Schaffers, H. (2012). Special Issue on Smart Cities and the Future Internet in Europe. *Journal of the Knowledge Economy*, *4*(2), 119–134. https://doi.org/10.1007/s13132-012-0083-x

201. Kontogianni, A., & Alepis, E. (2020). Smart tourism: State of the art and literature review for the last six years. *Array*, *6*, 100020. https://doi.org/10.1016/j.array.2020.100020

202. Koo, C., Gretzel, U., & Donnellan, B. (2021). Special issue on smart tourism cities. *Asia Pacific Journal of Tourism Research*, *26*(4), 349–351. https://doi.org/10.1080/10941665.2021.1898905

203. Koo, C., Gretzel, U., Hunter, W. C., & Chung, N. (2015). The Role of IT in Tourism. *Asia Pacific Journal of Information Systems*, *25*(1), 99–104. https://doi.org/10.14329/apjis.2015.25.1.099

204. Kozak, R., & Sezgin, E. (2013). Examining domestic transactions of incoming tourists with credit cards in Turkey. *Management : Journal of Contemporary Management Issues*, *18*(1), 23–43.

205. Kvale, S., & Brinkmann, S. (2015). *Interview: Det kvalitative forskningsinterview som håndværk* (3rd ed.). Hans Reitzels Forlag. http://hansreitzel.dk/Metode-og-videnskabsteori/Interview/9788741263779

206. Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

207. Lam, C., & McKercher, B. (2013). The tourism data gap: The utility of official tourism information for the hospitality and tourism industry. *Tourism Management Perspectives*, *6*, 82–94. https://doi.org/10.1016/j.tmp.2012.12.003

208. Lamsfus, C., Alzua-Sorzabal, A., Torres-Manzanera, E., Martín Del Canto, D., Ordoñez, D., & Vallejo, I. (2012). *Theoretical Framework for a Tourism Internet of Things: Smart Destination*. 15–21.

209. Lamsfus, C., Martín, D., Alzua-Sorzabal, A., & Torres-Manzanera, E. (2015). Smart Tourism Destinations: An Extended Conception of Smart Cities Focusing on Human Mobility. In I. Tussyadiah & A. Inversini (Eds.), *Information and Communication Technologies in Tourism 2015* (pp. 363–375). Springer. https://ideas.repec.org/h/spr/sprchp/978-3-319-14343-9_27.html

210. Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. META Group.

211. LaValle, S. (2009). *Business Analytics and Optimization for the Intelligent Enterprise*. IBM Institute for Business Value. https://silo.tips/download/business-analytics-and-optimization-for-the-intelligent-enterprise

212. LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010). Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, *52*(2), 21–31.

213. Lee, C.-H., Birch, D., Wu, C., Silva, D., Tsinalis, O., Li, Y., Yan, S., Ghanem, M., & Guo, Y. (2013). Building a generic platform for big sensor data application. *2013 IEEE International Conference on Big Data*, 94–102. https://doi.org/10.1109/BigData.2013.6691559

214. Li, D., & Yang, Y. (2017). GIS Monitoring of Traveler Flows Based on Big Data. In *Analytics in smart tourism design* (pp. 111–126). Springer International Publishing. https://doi.org/10.1007/978-3-319-44263-1_7

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

216. Liburd, J., & Edwards, Dr. D. (Eds.). (2018). *Collaboration for Sustainable Tourism Development* (1st ed.). Goodfellow Publishers. https://doi.org/10.23912/9781911635000-3879

217. Lim, C., Kim, K.-J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities*, *82*, 86–99. https://doi.org/10.1016/j.cities.2018.04.011

218. Living in EU. (n.d.). *Urban Data Platform*. Retrieved March 31, 2022, from https://living-in.eu/groups/solutions/urban-data-platform

219. Lopez de Avila, A. (2015). *Smart destinations: XXI century tourism*. 4–6.

220. López de Ávila Muñoz, A., & García Sánchez, S. (2013, May). Destinos turísticos inteligentes. *Harvard Deusto*, 58–76.

221. Lozano-Oyola, M., Blancas, F. J., González, M., & Caballero, R. (2012). Sustainable tourism indicators as planning tools in cultural destinations. *Ecological Indicators*, *18*, 659–675. https://doi.org/10.1016/j.ecolind.2012.01.014

222. Lu, W., & Stepchenkova, S. (2015). User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software. *Journal of Hospitality Marketing & Management*, *24*(2), 119–154. https://doi.org/10.1080/19368623.2014.907758

223. Magaldi, D., & Berler, M. (2020). Semi-structured Interviews. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 4825–4830). Springer International Publishing. https://doi.org/10.1007/978-3-319-24612-3_857

224. Malhotra, N. K. (2008). *Investigación de mercados* (5th ed.). Pearson Education. http://www.ebooks7-24.com/?il=3724

225. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. https://bigdatawg.nist.gov/pdf/MGI_big_data_full_report.pdf

226. Mariani, M., Baggio, R., Fuchs, M., & Höpken, W. (2018). Business intelligence and big data in hospitality and tourism: A systematic literature review. *International Journal of Contemporary Hospitality Management*, *30*(12), 3514–3554. https://doi.org/10.1108/IJCHM-07-2017-0461

227. Marine-Roig, E., & Anton Clavé, S. (2015). Tourism analytics with massive user-generated content: A case study of Barcelona. *Journal of Destination Marketing & Management*, *4*(3), 162–172. https://doi.org/10.1016/j.jdmm.2015.06.004

228. Mason, P., Augustyn, M., & Seakhoa-King, A. (2010). Exploratory study in tourism: Designing an initial, qualitative phase of sequenced, mixed methods research. *International Journal of Tourism Research*, *12*(5), 432–448. https://doi.org/10.1002/jtr.763

229. Masseno, M. D., & Santos, C. (2018). Smart Tourism Destinations Privacy Risks on Data Protection. *Revista Eletrônica Sapere Aude*, *1*(1), 125–149.

230. Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think* (p. 242). Houghton Mifflin Harcourt.

231. Mazanec, J., & Wöber, K. (2010). *Analysing International City Tourism* (2nd ed.). Springer. https://doi.org/10.1007/978-3-211-09416-7

232. McAfee, A., & Brynjolfsson, E. (2012, October 1). Big Data: The Management Revolution. *Harvard Business Review*, *90*(10), 60–68.

233. McFarland, D. A., & McFarland, H. R. (2015). Big Data and the danger of being precisely inaccurate. *Big Data & Society*, *2*(2). https://doi.org/10.1177/2053951715602495

234. Mehmetoglu, M. (2004). *Kvalitativ metode for merkantile fag* (1st ed.). Fagbokforlaget. https://www.akademika.no/kvalitativ-metode-merkantile-fag/mehmetoglu-mehmet/9788276749946

235. Meijer, A., & Rodríguez Bolívar, M. P. (2015). Governing the smart city: A review of the literature on smart urban governance. *International Review of Administrative Sciences*, *82*(2), 392–408. https://doi.org/10.1177/0020852314564308

236. Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A Big Data Analytics Method for Tourist Behaviour Analysis. *Information & Management*, *54*(6), 771–785. https://doi.org/10.1016/j.im.2016.11.011

237. Michalik, P., Štofa, J., & Zolotová, I. (2014). Concept definition for Big Data architecture in the education system. *2014 IEEE 12th International Symposium on*

Applied Machine Intelligence and Informatics (SAMI), 331–334. https://doi.org/10.1109/SAMI.2014.6822433

238. Miele, S., & Shockley, R. (2013). *Analytics: The real-world use of big data*. IBM Corporation.

239. Milano, C., Cheer, J. M., & Novelli, M. (Eds.). (2019). *Overtourism: Excesses, Discontents and Measures in Travel and Tourism*. CABI. https://www.emerald.com/insight/content/doi/10.1108/JTF-09-2019-088/full/html

240. Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook, 2nd ed* (pp. xiv, 338). Sage Publications, Inc.

241. Mill, R. C., & Morrison, A. M. (2012). *The tourism system* (Vol. 7). Kendall Hunt.

242. Minghetti, V., & Buhalis, D. (2010). Digital Divide in Tourism. *Journal of Travel Research*, *49*(3), 267–281. https://doi.org/10.1177/0047287509346843

243. Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

244. Morales-Urrutia, X., Morales-Urrutia, D., Simabaña-Taipe, L., & Ramírez, C. A. B. (2020). Smart tourism and the application of ICT: The contribution of digital tools. *Iberian Journal of Information Systems and Technologies*, *32*, 64–76.

245. Morrison, A. M., & Coca-Stefaniak, J. A. (2020). *Routledge Handbook of Tourism Cities*. Routledge. https://doi.org/10.4324/9780429244605

246. Mruck, K., & Breuer, F. (2003). Subjectivity and Reflexivity in Qualitative Research—The FQS Issues. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *4*(2), Article 2. https://doi.org/10.17169/fqs-4.2.696

247. Musa, S. (2016). Smart Cities—A Roadmap for Development. *Journal of Telecommunications System & Management*, *5*(3). https://doi.org/10.4172/2167-0919.1000144

248. Nam, T., & Pardo, T. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, 282–291. https://doi.org/10.1145/2037556.2037602

249. Ndou, V., & Beqiri, M. (2015). Introduction for the Special Issue on Big Data. *EJASA: Decision Support Systems and Services Evaluation*, *5*(1), 1–3. https://doi.org/10.1285/i2037-3627v5n1p1

250. Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). Current trends in Smart City initiatives: Some stylised facts. *Cities*, *38*, 25–36. https://doi.org/10.1016/j.cities.2013.12.010

251. Neuhofer, B., Buhalis, D., & Ladkin, A. (2012). Conceptualising technology enhanced destination experiences. *Journal of Destination Marketing & Management*, *1*(1), 36–46. https://doi.org/10.1016/j.jdmm.2012.08.001

252. Neuhofer, B., Buhalis, D., & Ladkin, A. (2015). Smart technologies for personalized experiences: A case study in the hospitality domain. *Electronic Markets*, *25*(3), 243–254. http://dx.doi.org/10.1007/s12525-015-0182-1

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

254. Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., & Agha, R. (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery (London, England)*, *78*, 185–193. https://doi.org/10.1016/j.ijsu.2020.04.018

255. Nutley, S., & Davies, H. T. O. (2000). Getting Research into Practice: Making a Reality of Evidence-Based Practice: Some Lessons from the Diffusion of Innovations. *Public Money & Management*, *20*(4), 35–42. https://doi.org/10.1111/1467-9302.00234

256. Oates, G. (2016, February 1). How the Rise of Smart Cities is Impacting Travel and Tourism. *Skift*. https://skift.com/2016/02/01/how-the-rise-of-smart-cities-is-impacting-travel-and-tourism/

257. Oficina de Ciudad Inteligente. (n.d.). *València al minut | Smart City València*. Retrieved April 28, 2022, from https://www.valencia.es/valenciaalminut/

258. Okwechime, E., Duncan, P., & Edgar, D. (2018). Big data and smart cities: A public sector organizational learning perspective. *Information Systems and E-Business Management*, *16*(3), 601–625. https://doi.org/10.1007/s10257-017-0344-0

259. Oussous, A., Benjelloun, F.-Z., Ait Lahcen, A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University - Computer and Information Sciences*, *30*(4), 431–448. https://doi.org/10.1016/j.jksuci.2017.06.001

260. Ozer, K., & Cetin, G. (2021). Utilizing Big Data in Smart Destination. In C. Cobanoglu & V. Della Corte (Eds.), *Proceedings of the Global Conference on*

Services Management (pp. 126–128). USF M3 Publishing. https://doi.org/10.5038/2641-5062-v2

261. Pathak, V., Jena, B., & Kalra, S. (2013). Qualitative research. *Perspectives in Clinical Research*, *4*(3), 192. https://doi.org/10.4103/2229-3485.115389

262. Patton, M. Q., & Cochran, M. (2002). *A guide to using qualitative research methodology*. Medecins Sans Frontieres.

263. Peek, G.-J., & Troxler, P. (2014). *City in Transition: Urban Open Innovation Environments as a Radical Innovation*. 1–10.

264. Pérez, F. S. (2015). Big Data. *Economía industrial*, *395*, 71–86.

265. Perrey, J., Spillecke, D., & Umblijs, A. (2013). Smart analytics: How marketing drives short-term and long-term growth. In *Marketing & Sales: Big Data, Analytics, and the Future of Marketing & Sales* (pp. 24–29). Mc Kinsey Quarterly.

266. Phellas, C., Bloch, A., & Seale, C. (2011). Structured methods: Interviews, questionnaires and observation. In C. Seale (Ed.), *Researching Society & Culture* (3rd ed., pp. 181–205). SAGE Publications Ltd.

267. Porter, M. E., & Heppelmann, J. E. (2014, November 1). How Smart, Connected Products Are Transforming Competition. *Harvard Business Review*, *92*(11), 64–88.

268. Power, D. J. (2014). Using 'Big Data' for analytics and decision support. *Journal of Decision Systems*, *23*(2), 222–228. https://doi.org/10.1080/12460125.2014.888848

269. Pramanik, M. I., Lau, R. Y. K., Demirkan, H., & Azad, Md. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems with Applications*, *87*, 370–383. https://doi.org/10.1016/j.eswa.2017.06.027

270. Presenza, A., Micera, R., Splendiani, S., & Del Chiappa, G. (2014). Stakeholder e-involvement and participatory tourism planning: Analysis of an Italian case study. *International Journal of Knowledge-Based Development*, *5*(3), 311–328. https://doi.org/10.1504/IJKBD.2014.065320

271. Pries, K. H., & Dunnigan, R. (2015). *Big Data Analytics: A Practical Guide for Managers*. Auerbach Publications. https://doi.org/10.1201/b18055

272. Proton Technologies AG. (n.d.). *General Data Protection Regulation (GDPR) Compliance Guidelines*. GDPR.Eu. Retrieved April 28, 2022, from https://gdpr.eu/

273. RapidMiner. (n.d.). *Advanced Analytics vs Business Intelligence*. Retrieved May 20, 2022, from https://rapidminer.com/glossary/advanced-analytics-vs-bi/

274. Red Española de Ciudades Inteligentes. (n.d.). *Objetivos*. Retrieved April 28, 2022, from https://reddeciudadesinteligentes.es/objetivos/

275. Rehman, A. A., & Alharthi, K. (2016). An Introduction to Research Paradigms. *International Journal of Educational Investigations*, *3*(8), 51–59.

276. Robinson, R. (2012, September 26). The new architecture of Smart Cities. *The Urban Technologist*. https://theurbantechnologist.com/2012/09/26/the-new-architecture-of-smart-cities/

277. Rodríguez-Molano, J. I., Contreras-Bravo, L. E., & López-Santana, E. R. (2018). Big Data Tools for Smart Cities. In Y. Tan, Y. Shi, & Q. Tang (Eds.), *Data Mining and Big Data* (Vol. 10943, pp. 649–658). Springer International Publishing. https://doi.org/10.1007/978-3-319-93803-5_61

278. Rotchanakitumnuai, S. (2017). Barriers to Bangkok as a Smart Destination with Internet of Things Technology. *Thammasat Review*, *20*(2), 1–17. https://doi.org/10.14456/TUREVIEW.2017.7

279. Ruiz Garcia, S., & Hernández García, Y. (2017). Impacto de las TIC en el sector turístico y su importancia. *Universidad & Ciencia*, *6*(3), 66–76.

280. Russom, P. (2011). Big Data Analytics. *TDWI Best Practices*. https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf

281. Saluveer, E., Raun, J., Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A., & Silm, S. (2020). Methodological framework for producing national tourism statistics from mobile positioning data. *Annals of Tourism Research*, *81*, 102895. https://doi.org/10.1016/j.annals.2020.102895

282. Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing & Health*, *23*(4), 334–340. https://doi.org/10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G

283. SAS. (n.d.). *Big Data Analytics: What it is and why it matters*. Retrieved April 15, 2022, from https://www.sas.com/en_us/insights/analytics/big-data-analytics.html

284. Saunders, M., Lewis, P., Thornhill, A., & Bristow, A. (2019). Understanding research philosophy and approaches to theory development. In *Research Methods for Business Students* (8th ed., pp. 128–171). Pearson Education.

285. Seagate UK. (2018, November 26). *Seagate Launches New Data-Readiness Index Revealing Impact Across Four Global Industries As 30 Percent Of Data Forecasted To Be Real-Time By 2025*. Seagate.Com.

https://www.seagate.com/gb/en/news/news-archive/seagate-launches-data-readiness-index-pr-master/

286. Secretaría de Estado de Turismo. (2012). *Plan Nacional e Integral de Turismo*. 124.

287. SEGITTUR. (n.d.). *Lines of action*. Retrieved April 28, 2022, from https://www.segittur.es/en/lines-of-action/

288. Shah, S., Horne, A., & Capellá, J. (2012, April 1). Good Data Won't Guarantee Good Decisions. *Harvard Business Review*, *90*, 23–25.

289. Shapiro, J. M. (2006). Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital. *The Review of Economics and Statistics*, *88*(2), 324–335. https://doi.org/10.1162/rest.88.2.324

290. Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, *191*, 97–112. https://doi.org/10.1016/j.ijpe.2017.06.006

291. Shoval, N., Isaacson, M., & Chhetri, P. (2014). GPS, Smartphones, and the Future of Tourism Research. In *The Wiley Blackwell Companion to Tourism* (pp. 251–261). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118474648.ch20

292. Sigala, M., Christou, E., & Gretzel, U. (2012). *Social media in travel, tourism and hospitality: Theory, practice and cases*. Ashgate Pub. http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9781409420927

293. Sigalat-Signes, E., Calvo-Palomares, R., Roig-Merino, B., & García-Adán, I. (2020). Transition towards a tourist innovation model: The smart tourism destination: Reality or territorial marketing? *Journal of Innovation & Knowledge*, *5*(2), 96–104. https://doi.org/10.1016/j.jik.2019.06.002

294. Smart City Office. (n.d.-a). Connecta VLCi: 194 smart buildings and municipal facilities. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/vlci/connecta-vlci-smart-buildings/

295. Smart City Office. (n.d.-b). GIS Portal: València City Council resources available to the citizen on a map. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/vlci/gis-portal/

296. Smart City Office. (n.d.-c). Smart City Office. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/smart-city-office/

297. Smart City Office. (n.d.-d). Smart Tourism. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/vlci/smart-tourism/

298. Smart City Office. (n.d.-e). València joins the European WiFi4EU network offering free and fast WiFi. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/vlci/wifi-4-eu/

299. Smart City Office. (n.d.-f). València Smart City real time information dashboard: València al minut. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/vlci/valencia-al-minut-city-dashboard/

300. Smart City Office. (n.d.-g). VLCi Platform. *València Smart City*. Retrieved April 28, 2022, from https://smartcity.valencia.es/vlci/vlci-platform/

301. Smart City Office. (2021, May 10). Why València is a leader in smart city KPIs. *València Smart City*. https://smartcity.valencia.es/why-valencia-is-leader-in-smart-city-kpis/

302. Smyth, R. (2008). NVivo (Software). In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (1st ed., pp. 562–563). Sage Publications.

303. Song, H., & Liu, H. (2017). Predicting Tourist Demand Using Big Data. In *Analytics in Smart Tourism Design* (pp. 13–29). Springer International Publishing. https://www.springerprofessional.de/predicting-tourist-demand-using-big-data/10864856

304. Soteriades, M. (2012). Tourism Destination Marketing: Approaches Improving Effectiveness and Efficiency. *Journal of Hospitality and Tourism Technology*, *3*(2), 107–120.

305. Stepchenkova, S. (2012). Content Analysis. In L. Dwyer, A. Gill, & N. Seetaram (Eds.), *Handbook of Research Methods in Tourism*. Edward Elgar Publishing. https://www.elgaronline.com/view/edcoll/9781781001288/9781781001288.00031.xml

306. Struijs, P., Braaksma, B., & Daas, P. J. (2014). Official statistics and Big Data. *Big Data & Society*, *1*(1), 2053951714538417. https://doi.org/10.1177/2053951714538417

307. Su, K., Li, J., & Fu, H. (2011). Smart City and the Applications. *Electronics, Communications and Control (ICECC)*, 1028–1031.

308. 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. https://doi.org/10.1109/ACCESS.2016.2529723

309. Swan, M. (2013). The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data*, *1*(2), 85–99. https://doi.org/10.1089/big.2012.0002

310. Tanggaard, L., & Brinkmann, S. (2015). Interviewet: Samtalen som forskningsmetode. In S. Brinkmann & L. Tanggaard (Eds.), *Kvalitative metoder: En grundbog* (2nd ed., pp. 29–53). Hans Reitzels Forlag.

311. Tapadinhas, J. (2014, June 10). *Business analytics from basics to value*. https://www.slideshare.net/sucesuminas/business-analytics-from-basics-to-value

312. The Economist. (2013, September 5). The multiplexed metropolis. *The Economist*. https://www.economist.com/briefing/2013/09/05/the-multiplexed-metropolis

313. Townsend, A. M. (2013). *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. W. W. Norton & Company.

314. Tracy, S. J. (2013). *Qualitative Research Methods*. Wiley-Blackwell.

315. Trencher, G., & Karvonen, A. (2019). Stretching "smart": Advancing health and well-being through the smart city agenda. *Local Environment*, *24*(7), 610–627. https://doi.org/10.1080/13549839.2017.1360264

316. Tribe, J. (2001). Research Paradigms and the Tourism Curriculum. *Journal of Travel Research*, *39*(4), 442–448. https://doi.org/10.1177/004728750103900411

317. Tu, Q., & Liu, A. (2014). *Framework of Smart Tourism Research and Related Progress in China*. 140–146.

318. Tussyadiah, I. (2013). Expectation of Travel Experiences with Wearable Computing Devices. In Z. Xiang & I. Tussyadiah (Eds.), *Information and Communication Technologies in Tourism 2014* (pp. 539–552). Springer. https://ideas.repec.org/h/spr/sprchp/978-3-319-03973-2_39.html

319. Tyan, I., Yagüe, M. I., & Guevara-Plaza, A. (2020). Blockchain Technology for Smart Tourism Destinations. *Sustainability*, *12*(22), 9715. https://doi.org/10.3390/su12229715

320. Ularu, E. G., Puican, F. C., Apostu, A., & Velicanu, M. (2012). Perspectives on Big Data and Big Data Analytics. *Database Systems Journal*, *3*(4), 3–14.

321. UNE. (n.d.). *Comité CTN 178 Ciudades inteligentes*. Retrieved April 28, 2022, from https://www.une.org/encuentra-tu-norma/comites-tecnicos-de-normalizacion/comite?c=CTN%20178

322. UNWTO. (n.d.-a). *Digital Transformation*. Retrieved April 5, 2022, from https://www.unwto.org/digital-transformation

323. UNWTO. (n.d.-b). *Glossary of tourism terms*. Retrieved March 9, 2022, from https://www.unwto.org/glossary-tourism-terms

324. UNWTO. (n.d.-c). *UN standards for measuring tourism*. Retrieved May 26, 2022, from https://www.unwto.org/standards/un-standards-for-measuring-tourism

325. UNWTO. (n.d.-d). *Why Tourism?* Retrieved March 14, 2022, from https://www.unwto.org/why-tourism

326. UNWTO. (2012). *Global Report on City Tourism*. UNWTO. https://doi.org/10.18111/9789284415300

327. UNWTO. (2019). *UNWTO Tourism Definitions*. https://www.e-unwto.org/doi/epdf/10.18111/9789284420858

328. UNWTO. (2020). *Supporting jobs and economies through travel & tourism. Call for action to mitigate the socio-economic impact of COVID-19 and accelerate recovery.* https://webunwto.s3.eu-west-1.amazonaws.com/s3fs-public/2020-04/COVID19_Recommendations_English_1.pdf

329. UNWTO & UNDP. (2017). *Tourism and the Sustainable Development Goals – Journey to 2030*. World Tourism Organization (UNWTO). https://doi.org/10.18111/9789284419401

330. Usher, R. (1996). A critique of the neglected epistemological assumptions of educational research. In D. Scott & R. Usher (Eds.), *Understanding Educational Research* (pp. 9–32). Routledge.

331. Vanolo, A. (2014). Smartmentality: The Smart City as Disciplinary Strategy. *Urban Studies, 51*(5), 883–898. https://doi.org/10.1177/0042098013494427

332. Verhoef, P., Kooge, E., & Walk, N. (2016). *Creating value with big data analytics: Making smarter marketing decisions* (p. 316). Routledge. https://doi.org/10.4324/9781315734750

333. Visit Valencia. (n.d.). *Why Visit València | Why to go to València*. Retrieved April 28, 2022, from https://www.visitvalencia.com/en/plan-your-trip-to-valencia/valencia/reason-visit-valencia

334. Visit Valencia. (2020). *Tourism Statistics 2020*. https://fundacion.visitvalencia.com/sites/default/files/media/downloadable-file/files/Folleto_Estadi%CC%81sticas_2020_v6_digital.pdf

335. Visit Valencia. (2022). *Viajeros en hoteles y apartamentos. Valencia*. https://app.powerbi.com/view?r=eyJrIjoiOTczMDIyNTYtMjk1Yi00MzNiLWFmYWQtZmZiNWJlYWJlNGJjIiwidCI6ImM2ODQ2NmVmLWFiOWItNGYwZS04ZTczLWViMjIwYWYyNTc5NiIsImMiOjl9

336. Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management*, *46*(C), 222–232. https://doi.org/10.1016/j.tourman.2014.07.003

337. Wall, G. (1996). Integrating Integrated Resorts. *Annals of Tourism Research*, *23*(3), 713–717. https://doi.org/10.1016/0160-7383(95)00091-7

338. Walt Disney World Resort. (n.d.). *MagicBands & Admission Cards*. Unlock the Magic with Your MagicBand or Card. Retrieved April 18, 2022, from https://www.disneyworld.eu/plan/my-disney-experience/bands-cards/

339. Wang, D., Li, X. (Robert), & Li, Y. (2013). China's "smart tourism destination" initiative: A taste of the service-dominant logic. *Journal of Destination Marketing & Management*, *2*(2), 59–61. https://doi.org/10.1016/j.jdmm.2013.05.004

340. Wang, D., Park, S., & Fesenmaier, D. R. (2012). The Role of Smartphones in Mediating the Touristic Experience. *Journal of Travel Research*, *51*(4), 371–387. https://doi.org/10.1177/0047287511426341

341. Wang, D., & Xiang, Z. (2012). The New Landscape of Travel: A Comprehensive Analysis of Smartphone Apps. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and Communication Technologies in Tourism 2012* (pp. 308–319). Springer. https://doi.org/10.1007/978-3-7091-1142-0_27

342. Wang, D., Xiang, Z., & Fesenmaier, D. R. (2014). Adapting to the mobile world: A model of smartphone use. *Annals of Tourism Research*, *48*, 11–26. https://doi.org/10.1016/j.annals.2014.04.008

343. Wells, W. D., & Lo Sciuto, L. A. (1966). Direct Observation of Purchasing Behavior. *Journal of Marketing Research*, *3*(3), 227–233. https://doi.org/10.2307/3149897

344. White, M. (2012). Digital workplaces: Vision and reality. *Business Information Review*, *29*(4), 205–214. https://doi.org/10.1177/0266382112470412

345. Williams, A. M., & Baláž, V. (2015). Tourism Risk and Uncertainty: Theoretical Reflections. *Journal of Travel Research*, *54*(3), 271–287. https://doi.org/10.1177/0047287514523334

346. Wise, N. (2018). Using ethnographic and participatory approaches in tourism and hospitality research. In *Handbook of Research Methods for Tourism and Hospitality*

Management (pp. 115–122). Edward Elgar Publishing Ltd. http://www.scopus.com/inward/record.url?scp=85075818844&partnerID=8YFLogxK

347. Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques* (3rd ed.). Morgan Kaufmann.

348. WTTC. (n.d.). *WTTC Big Data Report*. Retrieved May 26, 2022, from https://de.scribd.com/document/464466170/WTTC-Big-Data-Report-final-pdf

349. WTTC. (2019a). *Benchmarking Research Trends 2019*.

350. WTTC. (2019b). *Destination 2030: Global Cities' Readiness For Tourism Growth*. World Travel & Tourism Council. https://wttc.org/Portals/0/Documents/Reports/2019/Global%20Cities%20Readiness%20For%20Tourism%20Growth-Jun%202019.pdf?ver=2021-02-25-182751-627

351. WTTC. (2020, October 29). *174m Travel & Tourism jobs could be lost due to COVID-19 and travel restrictions, says WTTC*. https://wttc.org/News-Article/174m-Travel-&-Tourism-jobs-could-be-lost-due-to-COVID-19-and-travel-restrictions

352. Xia, Y., Zhang, L., & Liu, Y. (2016). Special issue on big data driven Intelligent Transportation Systems. *Neurocomputing*, *181*, 1–3. https://doi.org/10.1016/j.neucom.2015.10.121

353. Xiang, Z., & Fesenmaier, D. R. (Eds.). (2017a). *Analytics in Smart Tourism Design: Concepts and Methods*. Springer International Publishing. https://doi.org/10.1007/978-3-319-44263-1

354. Xiang, Z., & Fesenmaier, D. R. (2017b). Big Data Analytics, Tourism Design and Smart Tourism. In Z. Xiang & D. R. Fesenmaier (Eds.), *Analytics in Smart Tourism Design* (pp. 299–307). Springer International Publishing. https://doi.org/10.1007/978-3-319-44263-1_17

355. Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, *31*(2), 179–188. https://doi.org/10.1016/j.tourman.2009.02.016

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

357. Xiang, Z., Tussyadiah, I., & Buhalis, D. (2015). Smart destinations: Foundations, analytics, and applications. *Journal of Destination Marketing & Management*, *4*(3), 143–144. https://doi.org/10.1016/j.jdmm.2015.07.001

358. Yang, Y., Pan, B., & Song, H. (2014). Predicting Hotel Demand Using Destination Marketing Organization's Web Traffic Data. *Journal of Travel Research*, *53*(4), 433–447. https://doi.org/10.1177/0047287513500391

359. Yigitcanlar, T., & Kamruzzaman, Md. (2018). Does smart city policy lead to sustainability of cities? *Land Use Policy*, *73*, 49–58. https://doi.org/10.1016/j.landusepol.2018.01.034

360. Yigitcanlar, T., Kamruzzaman, Md., Buys, L., Ioppolo, G., Sabatini-Marques, J., da Costa, E. M., & Yun, J. J. (2018). Understanding 'smart cities': Intertwining development drivers with desired outcomes in a multidimensional framework. *Cities*, *81*, 145–160. https://doi.org/10.1016/j.cities.2018.04.003

361. Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Sage Publications.

362. Zhou, K. Z., & Li, C. B. (2012). How knowledge affects radical innovation: Knowledge base, market knowledge acquisition, and internal knowledge sharing. *Strategic Management Journal*, *33*(9), 1090–1102.

363. Zhu, H., Chen, E., Xiong, H., Yu, K., Cao, H., & Tian, J. (2015). Mining Mobile User Preferences for Personalized Context-Aware Recommendation. *ACM Transactions on Intelligent Systems and Technology*, *5*(4), 1–27. https://doi.org/10.1145/2532515

364. Zhu, W., Zhang, L., & Li, N. (2014). Challenges, Function Changing of Government and Enterprises in Chinese Smart Tourism. In Z. Xiang & I. Tussyadiah (Eds.), *Information and Communication Technologies in Tourism 2014* (p. 4). Springer.

365. Zhu, Y., & Xiong, Y. (2015). Defining Data Science. *ArXiv:1501.05039 [Cs]*. https://doi.org/10.48550/arXiv.1501.05039

366. Zupan Korže, S. (2019). How Smart Tourism Embrace Blockchains and Smart Contracts. *Journal of Innovative Business and Management*, *11*(2), 32–40. https://doi.org/10.32015/JIBM/2019-11-2-4

APPENDICES

Appendix 1: Informed consent to participate in a research study

Study Title: Challenges and Opportunities of Leveraging Big Data in Smart Tourism Destinations

Researcher: Ella Pommeranz

You are being invited to participate in a research study. Please read this consent form carefully as it will provide you with information on the research project and what you will need to do. Your participation is voluntary.

Purpose:

As technology continues to develop at a fast pace and thanks to advancements such as the introduction of the internet, more and more data is being generated and stored. There is a heavy dependence of the tourism and hospitality industry on innovation, technology and the use of data. The notion of being "smart" has been a buzzword in recent years and is now being applied to tourism development as well. The recent popularity of smart tourism has gained the attention of many due to its ability to provide quality information and insights. This research, therefore, seeks to investigate the potential obstacles and opportunities of smart destinations in using Big Data. Your participation in this study will help us to better understand this topic and what lessons can be learned from the European Capital of Smart Tourism 2022 - Valencia.

Procedures:

The interview will last for approximately 30-45 minutes and consists of two major parts. The first part will test the respondent’s knowledge of the researched concepts, namely Big Data and Smart Tourism Destinations. The second part of the interview will ask questions about the application and acceptance of Big Data and technology in tourism, and in Valencia specifically. In particular, the utilization of the implemented Big Data platform of the City of Valencia will be of interest. The interviewees will also be asked questions about their opinions and perceptions on the subject.

Audio Recording:

The interview will be audio recorded to enable the researcher to obtain responses that might not have been written down during the interview. Recordings will be deleted after transcription and data analysis when this study has been completed.

Privacy and Confidentiality:

Your signed consent form will be kept separate from your study data, and responses will not be linked to you.

Contact Information:

If you have any questions or concerns about this research, you may contact Ljubica Knežević Cvelbar at ljubica.knezevic@ef.uni-lj.si.

Consent Statement and Signature:

I have read this consent form and have had the opportunity to have my questions answered to my satisfaction. I voluntarily agree to participate in this study. I understand that a copy of this consent will be provided to me for future reference.

_____ _____

Participant Signature Date

Appendix 2: Final version of the interview guideline

1. Opportunities and challenges

- What are the main opportunities of Big Data for the tourism sector?
- How can Big Data help cope with different crises coming from the broader environment (safety and security, health, economic, demographic, climate etc.)?
- What are the main challenges related to Big Data implementation in the tourism industry?

2. Data availability and access

- What type of local/national data (= information generated and collected in the community of Valencia) does your organization have access to and use in its operations? (e.g. overnights, arrivals, purchasing behaviour, movement patterns of tourists)
    - Do you know the VLCi platform and how are you using it?
- What type of global data does your company have access to and use in its operations? (e.g. Mabrian, STR, UNWTO, ForwardKeys, etc.)
- Are you co-sharing your data/ dashboard with other organizations or companies at the destination level? Why (not)?
- What type of data will be necessary to have access to in the future?

3. Data analysis and processing

- How is Big Data used in your organization?
- What type of basic analysis are you using to analyze and learn from past events in order to be reactive? (e.g. reporting, KPIs, dashboards, scorecards)
- What type of advanced analysis are you using to analyze the current situation and to predict future events and trends in order to be proactive? (e.g. data mining, simulations, predictive modeling)

4. Infrastructure

- What type of infrastructure does your organization currently have that enables the analysis of Big Data? (e.g. cloud storage, internet connection, sensors)
- What infrastructure would be needed for an optimal use and analysis of Big Data (now and in the future)?

5. Data analysis skills

- How would you describe the skills that employees in your organization currently have in regard to Big Data analysis?
- What other skills would be needed for an optimal use and analysis of Big Data (now and in the future)?

6. Future outlook

- What are some key factors that businesses or destinations need to optimize the implementation and performance of Big Data?
- How do you think will the adoption of Big Data in the tourism sector evolve in the future?

Appendix 3: Interview transcriptions

Transcript 1 – Interview with Carlos Romero Dexeus (Director of Tourism Research, Development and Innovation at SEGITTUR)

Date: March 16th, 2022

Duration: 1h 37min

Ella Pommeranz: First of all, I would like to look at what data you are using, and you have access to. I divided these two into local and national data, and then we have global data. So, what type of local or national data does your organization have access to and use in its operations?

Carlos Romero Dexeus: Well, traditionally, when we talk about tourism data, we make the difference between data provided by official national statistics and the rest of the information available. So, from the point of view of the actual official statistics, we have all the information provided by suppliers' statistics, statistics of different kinds of accommodation. In the case of Spain, we have four major surveys. The hotels' survey, the camping survey, vacation homes and houses survey and the rural accommodation survey. We have those four major supply-side tourism surveys. We get a lot of information from those surveys. They have certain limitations of course. They only provide information on a monthly basis. They rely their confidence on the quality of the answers provided by the corporations, accommodation managers and owners, and staff involved in filling in of the questionnaires. But we get a lot of information that is available in the questionnaires and the information available on the website of the national statistics office in Spain. It is not only accommodation, it is not only about the number of nights. It is also about pricing, occupation and about different distribution channels and using those in accommodation establishments. Then we use information from the demand side, the information that we get directly from asking the travelers. We have three major surveys of this kind in Spain. We have the expenditure survey, it is a survey that the national statistics office does at the borders: at road borders harbour borders, and airport borders all over Spain. And with this survey, we collect information from visitors coming to Spain from abroad, but also for Spanish residents coming back to Spain through those borders. And we collect information about not only expenditures but also about activities, about preferences, about places visited, about means of transport, about the type of organization of the trip, whether it is an independent trip or a packaged tour. A lot of the information, which we call EGATUR (Encuesta de Gasto Turistico). The information is also available through our website. We have another survey that we call FRONTUR (Estadística de Movimientos Turísticos en Frontera), another one collected on our borders. We make the difference between those two surveys. Through FRONTUR, in which we used a very rigid questionnaire with only 6 questions, we try to estimate the universe of the total number of visitors arriving to Spain on a monthly basis. It

is a kind of census. We combine information from those very brief questionnaires that we launch in the different means of transport.

Together with administrative registers provided by the airport authority, trade authority, road authority, and harbour front. And we mixed up this information from the administrative registers about the number of vehicles, planes, occupation of those planes, number of trains, occupation of those trains. Together with the information from the surveys, we build up the total population of travelers arriving to Spain on a monthly basis. Well together with administrative registers and traditional surveys, in FRONTUR we combine information from manual counting on road border crossing points. So we or the national statistics office select a sample of days over the year, and they have people counting vehicles and the occupation of those vehicles. The combination of those three types of information (…) gets us the number of arrivals segregated by country of origin, region of destination, the purpose of the visit, the main type of accommodation used and type of organisation, if it is an independent or a packaged tour trip. We expand the information that we collect through EGATUR to the total population provided by FRONTUR. It is a very unique combination of surveys. I do not know about any other country in the world that uses a survey like FRONTUR. We have been doing so since 1998 more or less. I know because I was part of the team involved in the design of those two operations, EGATUR and FRONTUR in Spain in 1999 and 2000. This information of expenditure is also used for our central bank for buyers and payment estimations. We have another source of information to understand domestic tourism. It is a household survey that we do in Spain with around 10,000 households that we analyze, or the logistical manager. We keep asking always the same households about the travel pattern of those households, where they go, how they go to those places, how much they spent. Well, all the critical information to characterize the domestic tourism in Spain. Those two types of information from the supply and the demand side are used to provide the official data required by EUROSTAT (European Statistical Office). There is a kind of EU directive which is mandatory for all the country members to accomplish with certain information under a certain periodicity. We just use those sources to feed this information asked by EUROSTAT. Well, this is national information, the level of official statistics. Together with this information, we have access to other registers. Not only those provided by transport authorities but also information from the Ministry of Labor, we call it "social security". They have very detailed information about the number of workers in each company, in each sector, in each region in Spain. And you could use the list of tourism characteristic activities or the tourism satellite account, which is a kind of national accounts statistical tool to measure the impact of tourism on the economy. There, you have the whole list of characteristic activities of tourism. With this list, together with the Ministry of Labor in Spain, we could get the numbers of how many people are already currently working in tourism companies. Well, it is not the only source of information we have to measure tourism employment, but we also have an employment survey. Practically every country in Europe has something similar to this, which is what we use to estimate the unemployment rate. In this case, it is not an administrative register, like in the case of the information provided by the Ministry of Labor.

It is a survey based on a subgroup, but it is very reliable. But again, there are some methodological differences between the two sources of information. The results are not always equivalent. They have their differences, but it is very easy to use both of them as a source of contrast if you want to talk about the characteristics of these unemployed: male, female, temporary, permanent, average salaries, a lot of information. We have also another register of companies in Spain. We call it the DERVE. It is something like the national companies register. It is by the national statistical office also, it is a kind of census of all the companies that exist in Spain organized by different sectors. So again, we could extract from there the companies directly linked with the tourism sector. Directly linked means that we are only considering passenger transport companies, accommodation, restaurants, travel agencies, recreational activities in destinations and cultural activities. But together with those characteristic activities, you have connected activities, they are not specifically touristic activities, but they are also involved in the process of producing a tourism service.

Ella Pommeranz: You really have access to a lot of data from a lot of different sources, which is very interesting and also the fact that you are comparing two sets of data that you have and looking for different results. That is very impressive, I have to say.

Carlos Romero Dexeus: Yeah, well, the other comparisons that you could do, and that we do sometimes, is the comparisons with other countries. So for instance, I have estimated here the amount of tourism expenditure of Spanish residents when they travel abroad and I can compare those data with the data from Germany, the information they have of Spanish tourists arriving to Germany and how much they spent. So there again, you have possibilities to start comparing data with equivalent sources of information.

What else? We have other sources of information. We have information from our tax agency. And again, they have a lot of information, but we only have access to a small part of the information they have. We always try to convince them to let us access more information. For instance, they provide us with information about what is the amount of money we received thanks to tourism through the different types of taxes that we applied, or the hotels or restaurants, or different services, which is a good way to explain the importance of how and why tourism is so important in Spain. But also, they have access to other types of information, for instance, we are now trying to negotiate the possibility of accessing the electricity consumption of households in all apartments in Spain. It is difficult to have access to this information through the electric companies because they do not supply it to you depending on the region of Spain. You have one company or the big players, maybe they could share one year or one month of information but not on a regular basis. The tax agency in Spain, they have access to this information on the monthly consumption, or probably even on a daily basis. We are trying to have access on a monthly basis to understand much better what the level of seasonality in certain destinations is. It is true that hoteliers provide you with partial information. But with the apartments it is much more difficult, to see if those apartments have been occupied, or not. You also have a lot of illegal apartments in the market. So, all these phenomena through indirect indicators, such as electricity consumption

or water consumption or production of waste are three good ways of estimating the boarding capacity of the destination and also to estimate the seasonality of the tourism activity in certain regions.

We get a lot of information also from the airport authority, it is called AENA – the Spanish national airport authority. They provide information for frontier usage purposes, but also we get information from them from two major databases that they have. They have information about all the fights which arrive at a certain airport, certain day, at a certain hour by type of plane, company, and occupation, the number of passengers on those flights. But also, they have information about bookings of slots four, five or even six months in advance. This means that companies book their slots in every airport in the world six months in advance. They compete for those slots, which are swift periods of 15 minutes, which is the time to land or depart. But this is a very good indicator to project and to foresee the potential arrivals maybe of the next Saturday or the next holidays from the airport authority. It is just that you have to adjust. You have to keep these slots information updated on a monthly basis because they keep changing until the last day.

Of course, we also receive information from the police, the national borders, from all those flows of visitors from outside the Schengen area but also information from visa. We do not use this information from the border control or the police to estimate the total amount of foreign visitors that we receive on a daily basis. It is FRONTUR, that uses this information. But we use this information as a source of contrast with the information of FRONTUR.

Ella Pommeranz: And would you say in general that, like all of the apartments, administrative departments and ministries that you mentioned are willing to share the data? Cause you said some of them would rather not do it. Are there some complications or is it getting better? Like how do you perceive that?

Carlos Romero Dexeus: It is, yeah, well, we have been working with them since a long time ago, many years ago, and it is always a process that they see how relevant it is for the Ministry of Tourism and tourism policy to have this information. It is getting easier now with the digitalization of all those Ministries because you can have access through API and is easier to define the variables. In the old days, it was much more difficult. But now it is a question of explaining it properly. We address it as a business, so it is different from one Ministry to the other. SEGITTUR is a public company, we belong to the Ministry [of Industry, Trade and Tourism], but it is always a request that comes from the Ministry itself, so that makes it easier. The most difficult institution is the tax agency.

So those are more or less the national traditional sources of information that we use. Again, we have been using it, we are still analyzing and considering how useful they are or not. Other potential sources of information, not official ones, they already have a lot of the new areas of knowledge linked with the internet world, and all the information which is there, reviews, social networks, online travel agencies which have a lot of information about

searches, bookings, about occupation, about characteristics of the supply of accommodation available online, and those big platforms. In the digital world, you have to apply different technologies. You have to interact with different providers. You have to use language technologies, maybe to understand what they are saying about Spain in 60 different languages. You have to interact with developers, with channel managers of hotels, the ones who are managing the way the hotel commercializes their rooms on the online market. And then another one is everything related to interaction with big companies such as telecom companies, finance companies, banking, rental car companies, credit card companies. These kinds of big multinationals, the big companies with data about their clients. So you have all of those tools. In both cases, you have to be ready to pay for data, understanding that sometimes you are not going to be able to work always with the same provider of information. This could be a problem. We found out in the past that for instance, we decided one year to buy the information from the telecom company called Telefonica and the following year we bought the information from Vodafone. We then compared the information for both of them and they were very different. Because they process the information in different ways with different definitions. They lack standards. There are big black boxes behind major operators in the way they process the information, but you do not have access to raw data. So you receive already processed data. So they are expensive, they are not comparable with others. You cannot commit always with the same provider. So clearly there are plenty of difficulties. It happens with telecom companies, but it happens also when we buy data from credit card companies, from MasterCard or from Visa. Or it can happen when you buy information from banks.

Also, the problem is that any one of them have the vision of the whole picture, of the whole market, and we only have a partial vision. So it is not easy. Sometimes they consider this kind of business as a residual part of their business – their main business is telecom services, or to issue credit cards. And this is a kind of lateral door through which they sell all their services, which means that sometimes they will put the attention, they will put their resources, they will dedicate quality type of reassurances to provide reliable data. They think that it is just a question of putting the information in the market, but without understanding the kind of definitions we use in tourism, understanding our sector, understanding our needs.

So, there are some difficulties. We keep open with them the guiding process of refining, retuning, learning together, trying to teach them how to provide useful information. So they do not substitute the official source of information. They are complementary sources that you take into consideration at a certain geographical level in which maybe you do not have information from the national site. So for instance, through FRONTUR and EGATUR and all those surveys, I could reach a geographical level of region, what we call tourism areas, which is smaller than a region. But it is really difficult for some big cities. There are a lot of areas in Spain that are not properly covered us with those official statistics. We do not have enough big samples of information. So we usually get this information, alternative

information which is already on the market, not official information, to understand the tourism database of the small and local tourism phenomenon.

Ella Pommeranz: Are all these local companies willing to provide you with the data if there are no official statistics? Or how do you get the data from the regions where there is not enough information?

Carlos Romero Dexeus: The problems of the official statistics have to do with periodicity, I only have data on a monthly basis in most of the cases. Also, in terms of a geographical approach of this information. So it is difficult for me to analyze the data of last weekend in Marbella or in Palma de Mallorca because maybe the official data that they have does not provide me with information for this specific weekend. Or it does not provide me with information for a specific location. And it is in those cases that I have to go to the market or the big operators to buy information from them. So I try to offer this information that we buy at the national level to the local destination. A different question is what is the kind of relations that they have, from the national perspective with the regional organizations and local organizations to have access to certain information. Because it is true that there are regions in Spain that have their own systems to collect their own regional and efficient tourism statistics. And sometimes on the local level also they have their own surveys and ways of collecting information locally.

So, you have to plan always in the case of Spain on those three levels, so national, regional and local. Traditionally we do not use the information they produce at the regional or local level. We keep this with the national regions. What we do is to provide them with our data. So usually the flow goes from the national level to the regional or local. But it is very difficult for the local authorities, it is complex to articulate processes by which I could have access on a monthly basis to regional and local data. So for me, if I have to do a study of the tourism evolution of small tourism villages in Spain, it is easier to do it without asking those villages. Directly get in touch with a telecom company or a big bank to get information from them and exploit this information on a local level. […]

There are certain companies like Mabrian or ForwardKey which have found a business opportunity by integrating different sources of information, official and non-official information, and are providing their services to destination managers at the local level. And they do it very well. In some cases, we have also signed contracts with those intermediates. We could call them tourism info mediators. We could sign some contracts with them to understand better how to process certain information we are not familiarized with or to copy or get inspiration from their dashboards or maybe just to use their information for our own purposes.

Ella Pommeranz: Which ones are you using or are you buying the data from?

Carlos Romero Dexeus: Well, we have a contract with Mabrian, we have a very good relationship with them. In the past, we had also a commercial relation with ForwardKeys.

Now we will probably launch another contract with ForwardKeys again. On a permanent basis, we have an open conversation with everyone, no matter if we have a contract or not. A couple of weeks ago I had a meeting with Mabrian, trying to understand better the new set of indicators that they have developed to measure tourism sustainability on the local level. So we opened a discussion about measuring sustainability at the local level from the point of view of tourism analysis. We also try to help them in the process of exporting their services to other latitudes, to South America or to Asia. Both of them are also part of our members, collaborators of the study network of our Smart Tourism Destinations.

Ella Pommeranz: Well, it is great that you are collaborating and working together cause, in the end, everyone is going to benefit from that. Then, my next question would be: after you collected all of this data - you already said that you are distributing it then to the regional and local level - so how are you co-sharing the data and how do you promote that?

Carlos Romero Dexeus: Well, we collect the data, we process the data, we organize data in different databases and we publish the information through our website, through a place that we call data store. In the data store, you have access not only to the tables, to the dashboards, but also to the files themselves. So different formats and you could just download those files. We are now trying to have an API system, as we still do not have it. But we received a lot of requirements from local authorities interested in having automatic daily access to our data, but up until now we have started through our website.

Ella Pommeranz: Okay. And then what type of data do you think will be necessary to have access to in the future or how is it going to develop in the future?

Carlos Romero Dexeus: Something that we realized over the last couple of years is that sometimes the official statistics with a monthly periodicity are not enough because things keep changing too fast. So probably the two areas in which we have to work much more to sophisticate our statistical data production is (1) periodicity. I am not talking about real-time, but at least on a weekly basis or daily basis, hours of the day. We need much more updated information. I want to know what happened yesterday or what would happen tomorrow, and everything made with forecasts and with modeling. I think we have been spending a lot of resources and time in processing, collecting, analyz[ing] the past, but we still have a lot to do with understanding the future. I see different scenarios that play with those data. And also probably much more graduality (2), in terms of breakdown. Not only [data] in cities but maybe also in areas within those cities. So being able to understand the tourism flows dynamics with much more detail. I think that it would be also very interesting. Probably we will have to put more artificial intelligence within our systems, we do not have artificial intelligence. But we do have a lot of statistics. And we are statisticians most of us, or economists, but still, we need more AI.

Ella Pommeranz: And even Big Data can help exactly in that regard because like if you have sensors everywhere, you can get real-time data. And also, for specific regions, like you can manage the flows of tourists and everything.

Carlos Romero Dexeus: Yeah, Big Data could be about real-time information or daily basis information, that is also Big Data. Big Data is also about understanding what they [tourists] are talking about, understanding the conversations that are taking place online. Understanding conversations in different languages and getting useful insights from those conversations. It has nothing to do with IoT, it has nothing to do with real-time. It has to do with using those machines to process languages. There are certain destinations which have already started to implement and to deploy a lot of sensors. We have cities such as Santander, Barcelona, Malaga or Benidorm, that have different kinds of sensors to measure different things. But till now they have a limited capacity of analysis. They have increased a lot their capacity of collected information, much more information from different sources, IoT sensors, drones, cameras, different kinds of sensors. They have increased a lot their capacity to collect all that information, organize it, but they are still missing the capacity to analyze it.

Ella Pommeranz: What do you think is the reason for that? Is it because they do not have skilled workers or because they simply do not have enough capacity in their company?

Carlos Romero Dexeus: At the beginning, they thought that their problem was that they did not have sensors. Someone told them that the problem was that they did not have cameras to register everything. Someone told him that the problem was that they did not have drones. And then suddenly they realize that that part of the technological infrastructure was the easy part of the problem. The complex part is this: For what reason did you want to know for every single minute how many people there are at this beach? What kind of public service, what kind of decision do you take in real-time related to this information? Not at all, not at all. I mean, real-time information is used in the case of the emergency service, and it is used by the police, those two units (police, emergency services, hospitals), for them it is critical. It could be critical; death or life decisions depend on real-time information. From the point of view of the manager of a destination, maybe they are starting to understand that it was not so critical to have all those sensors providing real-time information because at the end of the day, they do not have the instruments to manage the flows in real-time. There is a lot of people here that are asking: Now what? What is the question you wanted to answer with this? “I wanted to know how many of them [tourists] are there right now”. Yeah. But this answer, what does it tell you? Okay, now we have this beach in Marbella at five o’clock in the evening, we have 153 people. One minute before you had two people less, and one minute after you have two more. For what? We are trying to push them to understand. Maybe there are destinations managing their destinations in a very good way, with a very good governance without so much technology. Of course, they do require technology, but once you understood for what you would use that information. So I think that the big problem is the human skills or the human resources. The most scarce or rare resource in a touristic

destination is someone who is specialized in data or statistics, trained from an ethical point of view. You could find people working at the destination level specialized in marketing, specialized in MICE tourism, but it is very difficult to find destinations with these specific skills. People need to be trained with statistical capabilities and analytical skills. And there we have a big problem in Spain. We are investing a lot in technology, we are investing a lot in making a good diagnosis, understanding where the problems are, but the tourism management of the destination is getting more and more sophisticated and complex. But they still have the same kind of instructors, the same kind of tools that they had 23 years ago, the same kind of people and specialists that they had 20 years ago. But everything has changed, everything is much more complicated, tourists require a different interaction on the digital channels.

Ella Pommeranz: When we are already speaking about skills. What skills do you think workers in tourism currently have? What skills do they need and what skills are needed in the future?

Carlos Romero Dexeus: Well probably, the tourism skills are too general. They have to be more specific, specific to certain tasks, to certain functions. It is true that the problem is also that some of those functions have not been established already. So on the one hand, you are identifying those new functions that have to be assumed by a destination. On the other hand, you are asking for profiles of tourism managers with skills ready to attend those new functions. I think a much more digital perspective on everything is also required. People with skills open to interact with any kind of technology, device, much more used to interact with the digital markets. Tourism students do not usually end their careers [education] with a good knowledge of digital tourism and the digital tourism landscape. What kind of actors, what kind of roles? What kind of big platforms, how prices are fixed, what kind of competitors, what kind of technological providers? So, I would say we do require people with much more expertise in the digital world, the digital tools, but also people prepared to be good analysts, understanding how to work with evidence and not with intuition, to work with data, to get insights from data, and a stronger background on statistics, I would suggest. Technology, data processing and analysis capabilities are very critical and connected. And we do not have these kinds of profiles working in our destinations.

Ella Pommeranz: Yeah, it is very tricky. Like what you just said, I very much agree with you because in my master's program, I think I had one class digging into the digital sphere of tourism, but not really in-depth. So in the end we did not really learn a lot about statistics or other stuff. And only because I am interested, I am learning more about it, but it is not really based in the curriculum and in the end that is going to be missing, right. Like in the future, when we are going to be managers or something similar, we do need those skills. So I think you are very right about that.

Carlos Romero Dexeus: I think it is just a problem of putting in place training programs and practices. Something that we enforce afterwards in specific training programs.

Ella Pommeranz: Actually, I have a question because I think you mentioned earlier you are teaching companies how to provide better data or how to get more insights from it. I think you said something along those lines. So, do you have a program for that from SEGITTUR or how does that work?

Carlos Romero Dexeus: Yeah, we do not organize that as a program itself. It is a kind of relationship that we have with those operators. So, once I buy data from Vodafone, we spent a couple of months working with Vodafone for them to understand what is a tourist, what is a set of visitors? What is a resident? How are they going to estimate travelers in transit, this kind of information, this kind of processing. They have to learn how to adapt the information they try to sell. They have to understand how to adapt to the needs of the tourism sector. There is a process of training, interaction. Verbal interaction with the destination. It is a non-organized trading or familiarization process you have to do with a company that sells the data for the first time to the tourism sector. But also with managers of destinations, we are eager to buy quickly information from those operators without understanding really what they are buying. But there are problems that they could have with those data. Or they realize that the data are similar to the data they have already. What is the reason for those differences? Since five years ago we have increased the amount of time we dedicate to this kind of exchange of information, explaining how to use those new sources of information, what are the differences, for what you could use them, what problems they have. Because data providers on the one hand, and destination managers on the other, they try to simplify everything. But the answer is never simple.

Ella Pommeranz: Yeah, it is not just like a yes or no answer. It is always something in between.

Carlos Romero Dexeus: It is not a definitive solution. It is not just that "I am going to close all the traditional surveys and I am just going to use the information from the telecom company." Well, maybe this is a mistake.

Ella Pommeranz: Okay. So, you are working with the data providers and you are telling them how they can provide more valuable data, like how they can understand it better. And then you also work with the destination managers to understand the data that they have better and how to work with. Is that correct?

Carlos Romero Dexeus: Yes.

Ella Pommeranz: Amazing. I would just like to jump back one step again. So, how would you say is Big Data used in your company or is it used at all?

Carlos Romero Dexeus: We use a lot of data. Some of the sources of information that we are using could be considered as Big Data sources. For instance, when we analyze all our online reviews of travelers who have visited Spain in the last year to understand what they think about their experience in Spain, we are using Big Data. But we are having access to

this information through a Big Data provider. We buy this information from a company called ReviewPro. So the "real" Big Data is made by this company. The only thing we do is buy these data. They are the ones having the language technology to get the data from the reviews. But it is true that we have a lot of information from reviews that we have to organize in a certain way and manner with SQL, cubes, databases. We have to use different programs to process all this information that we get from traditional official national statistics. We have to use Python or R2 [probabilistic programming system] or different technology languages that we use to program and to process and organize information, the databases. If you ask me if we have any source of real-time data, no we do not at all. If we always want to have access to non-structurized information, or this structured information that Big Data is able to process, we require the services of a third party. Of a technological party that is specialized in processing. If we want to process images or photos from Twitter of places in Spain, we need to hire the services of a company, able to process all those images. They are doing the real Big Data. Not us.

Ella Pommeranz: You are just outsourcing it then.

Carlos Romero Dexeus: Yeah. Also because it is a kind of technology which evolves very, very quickly. It is faster, cheaper and simpler to hire those services. If we consider Big Data from the point of view of capacity to analyze differences of the information of different nature, organize them. Yes, we do that.

Ella Pommeranz: It seems like you are doing a lot with the data, but then again, not everyone has the capacity to work with. I think it is more like a step-by-step process, right? Like you cannot have everything all at once. It is like the destinations you talked about, now they have the sensors, but what are they doing with it? I think there are some steps before that you should follow before you can work with real-time data. Okay, just to recap: what type of basic analysis are you using to analyze past events? So this is more like reporting, KPIs, dashboards. I think you are doing all of the things I have just mentioned, right?

Carlos Romero Dexeus: Yes. It is also the statistical analysis of average values, percentages, series analysis, temporal series analysis, a little bit of modeling also. We have a focused modeling in place to estimate the evolution of tourism arrivals. It has to do with traditional statistics. Then we use different tools to build up the graphs, Power BI from Microsoft or Excel or Tabloid or R2 - different systems that we use to build up our report.

Ella Pommeranz: Well, you already said you are also doing modeling, which is part of the advanced analysis, right? Are you also doing something like simulations or predictions of future events?

Carlos Romero Dexeus: Well, no simulations. This is something we are considering to do because we have seen what is happening every day now. One day there is a volcano, the other day there is war in Ukraine. I think we have to be ready to do these kinds of simulations

for a terrorist attack or whatever, but we still will have this incapability, but we are able to foresee or predict the tourism arrivals of the next three months.

Ella Pommeranz: Let us move on to the potential barriers of Big Data analysis. And one of the ones that I found very often while doing my research is actually the infrastructure. This could be the cloud storage, internet connection, everything you need for doing Big Data or just data analysis. So what type of infrastructure do most tourism companies or Smart Tourism Destinations have at the moment. And again, what do they have now? What would they need and what would be needed in the future?

Carlos Romero Dexeus: Most of the destinations, most of the tourism companies, they do not have Big Data infrastructure apart from big companies. I mean, in Spain 95% of these tourism companies are SMEs or micro-companies. They do not have anything related to Big Data. Of course, big hotel chains, online travel agencies, tour operators, big restaurant chains. They have, I would not say Big Data, but they have Big Data repositories, of information from their clients. They have customer relationship management systems that are very sophisticated. They have content management systems to manage their content on the portals and some tourism websites. They have channel managers to connect with intermediaries in the online channels, they have property management systems to manage their tourism businesses. Well, all those technological platforms or technological tools have their own data repositories, inventories, repositories of pricing, clients and so on. At a destination, very few have Big Data. They have providers of data, providers of insights, they externalized most of those services, because it is the right thing to do. Externalized most of it because they are expensive systems. Even the IoT. Nowadays, you implement IoT in a destination through a license. So you hire someone, you pay 5,000 euros a month and they manage 10,000 sensors within your destination. You contract a service, you are not buying the technological infrastructure and implementing it. So, most of the Big Data services, nowadays and in the future, are going to be hired, are going to be contracted through services providers. No matter if it is a Wi-Fi infrastructure or IoT or sensors infrastructure, or even with drones, you are not going to buy it. You are going to buy the services of a company that allows you to use their drones for a couple of days or a couple of hours every year. But at the same time that you are a subcontracted on those services, you are required to have at least some capability inside to define these relations with the third party. You require some expertise on the insights, on the questions, on the analysis. You require this capability that I was telling you that they do not have right now.

So we end up again with the problem of human capabilities. I would say that in 90% of the cases the data capabilities or infrastructure is externalized to the organization. No matter if it is a destination, public company or a private company. You have other companies in the market which are those intermediates of data who are the ones really investing in those infrastructures. There is another company called BO Price. It is a Spanish company, like Mabrian or ForwardKeys, that provides services worldwide. But then again, they are very good with Big Data on all the information they collect from online travel agencies on prices,

accommodations, all kinds of information. I would say if I have the technological infrastructure as a tourism company, probably by next year I would sell it. I would subcontract to a third party. It is cheaper and more efficient. It is more convenient. It is up to date. So, they are moving there all those big technological infrastructures, outside of the organizations.

Ella Pommeranz: Okay. And what type of technological infrastructures is subcontracted? You said Wi-Fi, sensors, drones.

Carlos Romero Dexeus: Everything. All the dashboards. For instance, in the case of the city of Valencia, they designed it to build up its own data infrastructure in the same way we have at SEGITTUR, our own data infrastructure in terms of databases, in terms of dashboards, in terms of processing the data we collect from different sources. They have decided to build up their smart tourism intelligence system. We hired the services of our consultancy to develop the whole infrastructure, but at the same time, they hired the services of ForwardKeys to provide them with data and to provide them with the dashboard that ForwardKeys has. So they combined the two worlds. Instead of licensing. Madrid city did the same thing with a different technology, with Tabloid. It is a different technology than the one they used in Valencia, but again, they combined it. They have their own dashboards, but the database information at the same time, they have a contract with Mabrian.

Ella Pommeranz: Now, I would talk more generally about like the opportunities and challenges of Big Data. So, what would you say are the main opportunities of Big Data in the tourism sector?

Carlos Romero Dexeus: I think Big Data is a great opportunity in any complex sector where you have many variables, many interdependent elements, many different solutions, many impacts involved. As for instance, it happens in meteorology, Big Data is very useful there because you have a lot of information, a lot of reacting relationships between temperature, humidity, landscape, different regions of the world. In the case of tourism, because we are in a very complex system of actors, stakeholders and interactions, I think Big Data could be very useful to understand properly and much better these kinds of relations between the offline and the online, the different actors of the online, the evolution of pricing, the evolution with different trends. So there it could be useful together and combined with AI - artificial intelligence.

I think it could simplify a lot the way that destination managers are taking their decisions. I think it would help a lot to understand the effects of the tourism activity. On the territory, positive and negative effects. I think Big Data could be very useful for modeling. I feel there is a huge need for modeling in tourism. Big Data is critical to have basic information for different variables. I think it is going to be around those areas.

Ella Pommeranz: And what do you think about personalization and efficiency?

Carlos Romero Dexeus: Well, from the point of view of the private companies yes, not from the point of view of the public manager. So we have like different perspectives to try to understand the effect of a public policy, try to understand the negative and positive effect of tourism policy, try to understand the capacity of a destination and trying to understand what are the limits of the tourism activity. From the point of view of the company, they want to maximize the profits. Big Data could be a very good tool to maximize profits, to optimize procedures, to reduce the number of people involved in those procedures. We see the back-office of the company, and the relation with the clients, of course, you have to understand and anticipate their preferences, to modulate those preferences, to personalize tailor-made services, to co-create with their clients, to anticipate the needs that not even their clients know that they have in the next couple of hours, to anticipate situations. To manage online marketing, Big Data is already used a lot in digital marketing based on algorithms. Plenty, plenty of opportunities.

Ella Pommeranz: One thing that, I think we all realized by now, is that crisis, like for example the COVID-19 pandemic, can hit us anytime. How would you say can Big Data help us to cope with these kinds of crises, not only health-related but also demographic, climate, safety and security. You already mentioned terrorist attacks. Could you please elaborate on that?

Carlos Romero Dexeus: I do not know. We had already a lot of Big Data and we could have known what happened. And so I think at a certain level Big Data could help us to improve certain processes. I hope it will help us to understand better what is next. The influence of Big Data is already there on our mobiles or many decisions that we are taking already as individuals when we decide to travel in a certain way to a certain place. We keep checking on Skyscanner for different flight fares. Or when we check traffic jams in our cities or when we are looking for a new apartment and we compare prices. I think the data would help us to be more resilient. Resilient to any kind of impact or event from the outside of our sectors.

Ella Pommeranz: What are the main challenges related to Big Data in tourism?

Carlos Romero Dexeus: It is not technology. It is not prices or costs. It has to do with human capabilities, the right people with the right attitude and skills, properly prepared. It has to do with understanding the nature of the different sources of information. Being able to work with analysis, to interpret the results of Big Data.

Ella Pommeranz: In your opinion, what are some key factors that destinations or businesses need in order to optimize the implementation of Big Data?

Carlos Romero Dexeus: I think one of the greatest challenges they have is to understand: Big Data for what? I think that is critical. And who is going to manage it, who is going to get the insights? Who is going to take the decisions based on these new tools, these data? I think the challenge is to build a strategy for Big Data. Thinking about what sources with

what periodicity, at what cost, what is the cost of those sources of information, who is going to manage it? With what technological provider? The kind of partner you choose is also critical. The big challenge is about those issues.

Ella Pommeranz: Okay, great. And then last but not least, how do you think will the adoption of Big Data in the tourism sector look like? Do you think more tourism destinations are going to implement it or not really? Do you think it is just a hype?

Carlos Romero Dexeus: I think they will go with Big Data for sure. We probably will see a Big Data approach at different speeds. The big challenge of the tourism sector related to Big Data has to do with what I already told you. That type of data will be uploaded by most of the operators but at different speeds. It will be once the big players, the big companies, the big destinations, they would move very, very fast. The rest will lag a little bit behind. It will take them more time to join this process as well. Probably we will see more sophisticated, personalized services in real-time supporting us when we visit certain destinations. Big Data will be around in the future.

Ella Pommeranz: Well, thank you very much.

Transcript 2 – Interview with Daniel Pons Mallol (CEO at Bioparc Valencia)

Date: March 25th, 2022

Duration: 32 min

Ella Pommeranz: So, I wanted to look at what data you are using at Bioparc and I divided the data into the local or national data, so the data collected in Valencia, and then also global data. So, what type of local or national data does your organization have access to and use in its operations?

Daniel Pons Mallol: So, mostly what we use is our own data, so we have a very generic view of our type of visitors and we get that from when you purchase the tickets we always ask for the origin. We know the number of passes that you sell per ticket, so we estimate more or less the average size of the group that visits us, we know where they come from because we ask them. We know more or less the profile of the people that visit our website etc., so we have a pretty decent view of our customers. In the past, we have been able to correlate that number of visitors we received with the number of visitors that Valencia city received and we were doing some analysis and comparisons with that. We can see if we are following the trend of the city, if we are improving the trend of the city, if we are below the trend of the city. So, we measure these kinds of things. […] We did an analysis some years ago about the correlation between the average number of stays and our number of visitors,

etc, but in the end, it does not really affect our operations that much, because we are a park, we are open from ten to six or seven or eight, and we get as many visitors as we can. The impact on our operations is more or less the scheduling of activities and the number of personnel we have to attend to the visitors, these are the only things we have to take care of.

Ella Pommeranz: Okay, are you also analyzing the purchasing behavior of the visitors, how much they spend, where we spend it?

Daniel Pons Mallol: Well, we know how much they spend on us, we know how many days in advance they purchase [the tickets], we know how much they spend in our restaurants and shops yes.

Ella Pommeranz: And then, have you heard of the VLCi platform from the DMO of Valencia?

Daniel Pons Mallol: Well I am the CEO, I do not live in Valencia, we run four parks. Maybe my colleagues in Valencia have heard of it, I did not. So, I checked it out this morning.

Ella Pommeranz: Okay well that is also a very good answer because I feel like maybe not a lot of companies actually know the platform that Valencia is providing. So that is also one of the reasons why I am doing this research. Because in theory, I think it might be a very good idea but people have to actually use it.

Daniel Pons Mallol: Yeah and we do not use it.

Ella Pommeranz: So, what type of global data does your company have access to, or are you purchasing any data from other data providers?

Daniel Pons Mallol: No, no, not at all. Why would I need that data?

Ella Pommeranz: Well, I think that is really interesting for like tourism destination planning. And then the data that you are collecting, are you co-sharing that with someone else, like are you sharing it with the DMO [Visit Valencia] so they can plan better or are you not doing that at all?

Daniel Pons Mallol: Not really. Once per year we more or less state the number of visitors we had, the type of visitors, how many are international, how many are national. But it is true that in the last few years we have not done it because of COVID etc. We do some basic sharing. We do not use data that much. I mean we use it a lot internally to understand the performance of our business. And, as I said before, we have a product which is in a place that we cannot move, we can open the hours that we can open [depending on the sunlight]. So in the end, that is the only thing I can do.

Ella Pommeranz: Okay and are you collaborating with other tourism businesses to share the data?

Daniel Pons Mallol: No.

Ella Pommeranz: What type of data do you think will be necessary to have access to in the future?

Daniel Pons Mallol: Well, actually, I think, the one thing you mentioned is that knowing the tickets, the flight reservations and the train reservation and the hotel reservations in advance is always good. And in the end, we have that information because you talk with the sector. The sector is not very big, so you know hotel owners, you know people that work in the airport, etc. And you have an idea of "I have a lot of reservations" or "I do not have a lot of reservations", but that kind of information will be interesting to have.

Ella Pommeranz: Yeah in order to estimate the demand and when people are going to come.

Daniel Pons Mallol: Yeah, exactly. It is also true that lately, the last two years, with all the changes and all the different things that happen, it is very difficult to figure out what is going to happen, I mean Easter is in two weeks and I do not even know if this is going to be a good Easter. I do not even know if it is going to be good weather or bad weather. I mean for me the ideal information would be to know the weather.

Ella Pommeranz: So are you using the weather data?

Daniel Pons Mallol: Well, I use it and I check it, and I know what is going to happen, but if you ask me, I would love to know if on Easter it is going to rain or not.

Ella Pommeranz: Good, then we can move on to how you use all the data that you collected. So, in general, would you say that Big Data is used in your company or maybe not at all, or are you using basic data analysis?

Daniel Pons Mallol: Define Big Data. How big does the data sample have to be to be [considered] "big"?

Ella Pommeranz: I think that is the main issue is that there is not one number that I can tell you. It is more like a very large amount of data that you cannot store in just one place, so you need to use a cloud to use that and then analyze it.

Daniel Pons Mallol: I think from my understanding of Big Data we do not use it, we use data mining. Because we do not use database technology that is associated with Big Data. We use a standard database technology, SQL, traditional database manner and traditional data mining operation. That is what we do. On that, we are pretty good.

Ella Pommeranz: Okay, are you also using KPIs to analyze the development?

Daniel Pons Mallol: Yes.

Ella Pommeranz: So in that regard, are you using any type of advanced analysis, like as you said data mining, or in terms of simulations and predicting the future demand or not yet?

Daniel Pons Mallol: No, we do not do any predictions. We do simulations, of course, and we do a lot of data mining for performance review, but we do not do predictions.

Ella Pommeranz: And what kind of simulations are you doing?

Daniel Pons Mallol: Well, basically we are trying to simulate how different marketing approaches are going to work. If we are thinking about changing a price or if we are thinking about offering a new product or a new pass or something, we try to figure out what is going to happen.

Ella Pommeranz: Ok, so while I was doing my research, I came across a couple of barriers when implementing Big Data analysis. And I think one of the biggest ones I found is the infrastructure. Of course, you need a good Internet connection and you need cloud storage, maybe sensors depending on what you are doing. So what type of infrastructure does Bioparc Valencia currently have that enables the analysis of Big Data or just data?

Daniel Pons Mallol: Well, I mean we have everything on a data center. […] Basically, we have our server in a data center that it is managed by a third [party], and we use, as I said, relational database technology to do all that analysis. We have a pretty good Internet connection with several backups. Actually, all our data is currently in the cloud, we have no or only very little data is kept locally because we decided to move everything else. And the sensors, what we are doing is that we use a lot of sensors to check the status of our systems. But I mean when I talk about the systems then I talk about all the life support systems which are all the systems you have to maintain the aquariums and all the different enclosures in operation: heat, temperature, humidity, water levels, etc.

Ella Pommeranz: But you are not tracking the visitors?

Daniel Pons Mallol: No, no, because everybody tries to sell us to sell that kind of technology, but honestly I find it useless because I see very well where they go.

Ella Pommeranz: Yeah, that is also the thing because I think a lot of tourism businesses or like businesses in general, they would like this technology, but then, once they have the data it is always the question for what do you want this data.

Daniel Pons Mallol: Or they discovered that they already have the data. You already know "in this area I always have a lot of people and in this one, nobody ever comes".

Ella Pommeranz: Yeah and now you have the proof for that, that is true. Okay, so what type of infrastructure do you think would you need in the future for an optimal use of Big Data or even right now already?

Daniel Pons Mallol: I think for what we need we are pretty good. I do not think we need to know much more.

Ella Pommeranz: And another big barrier that I found was the skills of the employees in terms of analyzing the data. How would you describe the skills that your employees currently have?

Daniel Pons Mallol: I think our team, particularly marketing and technology, they have a pretty good understanding of what they are doing and what they are analyzing. We do not do Big Data, we do data mining. We have a pretty clear idea and we know how to use it and we know how to strike the information we want.

Ella Pommeranz: So, would you also say that the skills you have are sufficient like you would not need to improve them in the future? Or is there something that you would like to train?

Daniel Pons Mallol: There is always a need for improvement, there is always a need for learning more. But as I said before, I am not sure we need to do any jump to Big Data.

Ella Pommeranz: Okay, then let us talk more general about Big Data. So, what would you say are the main opportunities of Big Data for the tourism sector, or do you think there are any?

Daniel Pons Mallol: I mean, if I wanted to do something, I would, as I said before, try to predict the number of tourists you are going to receive in a particular season. For that you need to correlate it not only to the actual bookings because that is very easy, you have the booking so you count them, you should be able to correlate those bookings with external events. Those events are those that force the tourists to take a decision in one direction or another. For example, there is a war in Ukraine. Now everybody that will normally go to the East, will not go to the East. They will go somewhere else. They are going to the Western or they are going to Spain. I mean, that is the real value of Big Data that we will be able to do predictions, not only on the specific data set or data sample but correlating with other data samples. And it ensures that those data samples are those that induce the decisions in the future. I mean, knowing the number of bookings is cool but for that, I only need to call the hotels. “How many bookings do you have?” and they will tell you. I mean the bookings go up and down, they change. The point is how are people going to take the decision, why are they going to book in Valencia and not in the North. For example, if there is a lot of news about rain in Valencia, is that impacting the bookings in summer? Is it not going to impact it? How do I know? And that is the value of Big Data now. Is that possible? I really do not know, but the key point is not how many people are going to come. It is, what are the factors that are making people take a decision to go to one destination or another.

Ella Pommeranz: So, to sum up, you would say that Big Data enables better decision-making right. And then, what do you think about being able to react quicker to events like the war in Ukraine. Do you think that Big Data can help with that?

Daniel Pons Mallol: No, I think that the war in Ukraine is so big that is very evident that we are going to have more tourists this year.

Ella Pommeranz: Or any other event, just in general, do you think Big Data enables us to react more quickly to changes? Or if you know in real-time "Okay, this is going to happen", maybe you can adjust marketing.

Daniel Pons Mallol: I am not sure. As I said, what I think Big Data can add is measuring the impact of those events that are not huge. Because events that are huge are so easy to predict. I mean if in the case of Spain, if there are wars in Northern Africa all the Europeans are going to Tunisia or that area for holidays, they will not go. They look for going to Spain because it is in Europe but it is more expensive. This is an easy event to measure. Okay, but the smaller events like the impact of news in the long term, the impact of the weather and climate change, all these kinds of things, that is something that Big Data can prove a value because it can go beyond whatever an average person that walks in the market can predict easily. And I am not sure if people are thinking in that direction and I am not sure if it is possible and I am not really sure what is the real impact because if they tell me there are going to be very few visitors this year in Valencia, what do I do? Do I invest more to get them all, or because only a few are going to come do I invest less so I minimize my losses. Because publicity is a cost. So if there are going to be 1 million tourists then maybe I do not need to invest anything, because they will all come. It is not clear what is the impact of having that data in your decision-making. But for sure, if Big Data has to provide something, it is something that you cannot predict easily. And now, I do not see it.

Ella Pommeranz: I think what you said goes somewhere along the line of the simulations. You have to kind of like estimate: So if I do this marketing campaign, if I change prices, like what is going to happen.

Daniel Pons Mallol: Exactly, but that is another approach. I mean I could try to figure out how my marketing campaigns work in different years to see how many people are coming. But the fact of people coming is not related to one variable. It is marketing, it is the weather, the number of people in the city, it is the price. It is hundreds of things. And that is where Big Data could help.

Ella Pommeranz: Yes, for customer insights.

Daniel Pons Mallol: But it has to be extremely detailed, I mean it is really Big Data.

Ella Pommeranz: True. Actually, I do have a question about something we talked about earlier because you were not aware of the platform right, and they also smart tourism in

general. So I was just wondering if Visit Valencia or the Valencian government approached you and informed you about what they are doing, also about the smart tourism award.

Daniel Pons Mallol: I am sure they did.

Ella Pommeranz: They did, okay.

Daniel Pons Mallol: It is just that we have not figured out yet how to combine everything.

Ella Pommeranz: Okay yeah I mean it is a process. It is not that easy, right. Like we are doing it one day, and then the next day it works. But okay good that they reached out to you. Okay, and the next question is: what I think, as we all know by now, different crises like the COVID-19 pandemic can hit us anytime. And I was just wondering if you think that Big Data can help us to cope with these kinds of crises, like not only health crises, but also in terms of safety and security, economy, demographic crises or climate change?

Daniel Pons Mallol: I do not know. That is such an open question I have no idea. Technology, in general, can help us to cope with crises, yes. Big Data, in particular, it depends on the type of crisis.

Ella Pommeranz: Like what type of crisis, do you think it will help us with?

Daniel Pons Mallol: Honestly, I do not know, I do not have an answer.

Ella Pommeranz: That is alright, no worries. Okay, then on to the next question: what are the main challenges related to Big Data implementation in the tourism industry?

Daniel Pons Mallol: Well, I think the big challenge is to really understand the added value. I am not a hundred percent sure how we can use it. I mean, as we said, we could use it if you are able to correlate so many different factors but I am not sure how you take what the Big Data tells you, how you present it in your decision process. I think there is still a lot of work to be done and it is very important to understand who, what is the added value and who captures that value meaning. If you want to do a Big Data project, there is a lot of investment to be done by somebody, there is a lot of work to be done in development, in the data analysis, data scientists, engineers, etc, and those people need to be paid. So the question is: how do you generate all that value, how does the person or the entity that invests in that effort, get paid? In order to get paid one has to generate more value because otherwise, nobody will pay him and that is for me a business model problem. I mean I am not sure how somebody is going to make money with this. I can understand perfectly, I could make money if somebody comes and tells me “Listen, we have analyzed your campaign strategies over the last ten years, we have correlated it with the weather, the number of tourists, etc., and we discovered that this type of campaign works better”. That is great for me, but I mean to find that out you need to make such a huge investment. But I am not sure who is going to make it.

Ella Pommeranz: So who do you think, from what party should come that investment, from the government, the Spanish government or the Valencian government?

Daniel Pons Mallol: Well, I think if platforms like VLCi provide a lot of information that is generic information about the city, like the bookings etc. Maybe one private sector could decide to try to use all that data and correlate that with internal data. So I think it is a mix. I think there is a lot of data that needs to be available. It has to be placed available most probably by the municipality, and I think the VLCi [platform] is trying to do that. Maybe they need to present a lot more data, but then it is up to each private company to try to figure out how they use the data and implement it during the decision process.

Ella Pommeranz: So from what I have read this is exactly what they are trying to do. So, they are collecting a lot of data from the census or sensors in the city, then they combine this data, analyze it and publish it on that platform. And, apparently, they also do reports and send them out for free to tourism businesses […]. This is their main idea.

Daniel Pons Mallol: That is going in the right direction. But I have not checked the data yet.

Ella Pommeranz: Okay, so then we are coming to the last section, which is about the future of Big Data. In your opinion, what are some of the key factors that businesses need to optimize the implementation and performance of Big Data?

Daniel Pons Mallol: Well, I mean if we had to do a Big Data project, we would need for sure external help because we do not have experience in Big Data. We would know how to ask the questions, so what questions we want to know. But somebody else would have to do the analysis for us, that is for sure. And I do not think that a company like us should have a Big Data specialist, we should probably get training on what Big Data can do for us to try to understand the right questions and the questions are possible to answer. But it should be a third-party consultant or something that is able to consolidate all the data and give you the answer.

Ella Pommeranz: And do you think in general, tourism businesses do not need Big Data at all? Like it is more on the destination level?

Daniel Pons Mallol: No, I do not think that we do not need it. I think that the destination has to gather the data. And then the entities that are in that destination could use the data, but most of us are not really aware of what we could do with that data and I do not really know if we have the right data or not. It is a very tough analysis because a destination is also a complicated concept because in Valencia a lot of our visitors are not from Valencia city, they are from outside Valencia. So it is difficult to correlate. But I think, and I have to thank you for this, after this conversation, I think there is an opportunity, if the data is [already] gathered to optimize some of our processes.

Ella Pommeranz: Well, I hope, you are going to reach out to the platform, look at it, and maybe it is going to help you to improve your operations in the future, who knows. If that works, I would be very happy. Then I reached my goal. Okay, so, then the last question is, how do you think will the adoption of Big Data and the tourism sector evolve in the future?

Daniel Pons Mallol: I really do not know. I think that right now in the environment we have, we are all companies, we are trying to focus on the more immediate problems. I think we are in an environment in which long-term thinking has been stopped somehow because everything is changing so fast and in so weird directions. I think that if the data, and that is what I am realizing in this meeting, if the data is available somewhere because somebody gathers a massive amount of data, then I think there is an opportunity for tourism companies to improve their operations by asking the right questions to these data and crossing it with our own internal data. That, I think, is the way to do it. If I have this big set of data about Valencia: people, climate, prices, whatever. If I have it, I may be able to cross-compare it or try to correlate it with my own internal data and may be able to figure out some improvements.

Ella Pommeranz: Great, that sounds very good.

Daniel Pons Mallol: I think that this VLCi is pretty interesting.

Ella Pommeranz: Yeah, I think it is a great idea and that is why, again, I chose Valencia, because I do not think a lot of destinations have this kind of open data platform yet, and I want to see like if it works and then also what over destinations can learn from that. Okay, that was it. Thank you very much

---

Transcript 3 – Interview with Joan Carles Cambrils i Camarena (General Fundació Visit Valencia)

Date: April 1st, 2022

Duration: 1h 25min

Ella Pommeranz: So, first of all, I wanted to look at what data you are using and have access to. So, I divided these into local and national data - so data from Valencia - and then also global data. What type of local or national data does your organization, so Visit Valencia, have access to and use in its operations?

Joan-Carles Cambrils: Okay. So, I was thinking that the best idea at this moment is to show you some pages where you will see directly which type of information we deliver for our partners. At this moment, we deliver it open to everybody, even to you. We do not charge any fee for accessing our data. According to the needs of your research, I think there is evidence that we can not manage because we are a DMO, a destination management organization, all the city data that is displayed on other websites, [such as] in the city hall website. So, we should have a look over the city hall website and then, later on, we come to our website. Okay. So, you will understand better after this that when you are in a big city like Valencia or compared to Vienna or Hamburg or Ljubljana, there is no silo or two separate fields when you try to manage the city or manage the tourists. I mean, in these cities in Europe, when you receive tourists or visitors, they have total contact with locals. It is not the same if you choose to have a holiday in a beach area or the countryside or the mountains. The management of flows and the management of visitors there can be considered a quick and clear, separate issue for local tourism authorities or local governments. So, you should know that in cities, the management of citizens' needs is mixed with visitors' needs. For example, traffic control, parking areas, pollution, et cetera.

And for that reason, if you want to do further research in cities in Europe you should consider having a look over local statistics and local analyses coming from Big Data or conventional sources or official registrations and databases. Because here you can separate the areas where both types of people, both types of humans, live together. One lives permanently, and the other one temporarily. But they use more or less the same in public areas, they are sharing the public goods. In this situation, you need to look at where the sources of information are that can be used for tourists because they are deliberate for residents.

And so, I will share [my] screen. […] This website is called Valencia Al Minut. Unfortunately, you have [the information] just in the local language Valenciano or in the Spanish language. […] These are some of the display resources that Valencia city hall offers us and to everybody. They have a big platform that is called Valencia Smart City, Valencia Ciudad Inteligente, as you can see with my mouse there. You have specific information about Coronavirus. Well, you can go to check the ISO, the level at which Valencia is qualified and certified. Here, there are some other data of what we call "open government", which is about the local government, and the Geoportal which is a geo-reference database. Quite interesting. Geoportal is working over a geographic database. Okay. So here you have some data about the digital agenda. This is the street maps [showing] noise pollution and atmospheric pollution. Some other information about education, sports, where you can go for activities. Some other data about urban ecology. These are our *fallas*, our main local fiestas we have with those monuments. Some other open data, geodata, historical data of Valencia, information about gardening in Valencia, and a noise map. So, there are some stations measuring the noise in different places of the city. That is a good indicator of pollution, noise pollution. You can see here these red places. So, it is a quite interesting tool and compared to other cities all over the world, Valencia is quite well considered with these

resources. That is what I heard from officials from the city hall. That is because Valencia has two important levels in the smart city platforms. This ISO level is delivered following about a hundred indicators or KPIs, and there is another well-considered level in digital platforms delivered by ITU, that is the international telecommunications union. […[It is an institution well-considered all over the world because they deliver the standards for telecommunications and for the smart city. And Valencia has a platinum level, which is the highest level. I remember that for this certification, there are about 90 questions or 90 requirements. Maybe you got 75 or 70 and you go to the platinum level.

So we are having here, for example, public works in different places in the cities, so problems for traffic because in this avenue and others are works. And this is something you can check here in what is called "occupation of public ways". If you use this button, you see the traffic in real-time according to different cameras. And you can see that the green here is fine. Here, there is a small problem. And I think they get this information through public cameras that are in the city. […] You have here the urban bicycle paths. We are a best practice. And we are invited by many cities in Europe to talk about our bicycle path network, as you can see, it is all these red and blue [paths]. I do not know what the difference is now between red and blue, but all these are bicycle paths. […] If you went back with this button, you can see the EMT – Empresa Municipal de Transportes. This is the bus network. […] I think they should deliver you information about where is the bus at each moment. And we are watching all these on the same site, you know, this is powerful. Because if you download the municipal transport buses app, like many apps in Europe, you can check there in real-time the bus. But this is general information displayed on our website. Here is some meteorological weather information. I think they give you information about the temperature and humidity in different stations all over the city. Okay. We go to the next [one]: air quality. Well, these are stations that are measuring the air quality in the city. If it is green, this is in the surroundings, in the belts of the city, it is much better than this kind of blue. If they overpass, if they do not fit with the standards, it is a dangerous place. It will be displayed in red. So, this is objective and independent of political influences. So, it is something that all citizens can see. So, interesting information. And we have parking. In many of these parking [spots] they calculate the free parking or the busy parking in the area by cameras.

Valencia has a public network for bicycles that you can rent. In each city there is a different system, in this one you download an app, you [upload] your bank details and they charge you for each minute you are using a bike. So, you can get the bike in one of these places and leave it in another of one of these places. That is working quite well in Valencia but it is not a free bike use. That system was working here in Valencia, but it is not well-defined yet. But with this system, you can see where you can get a bike and if there are free bikes or there are no free bikes for you. [Here you can see] available bicycles at this time. 13 available bicycles. 12 places where you can leave your bicycle and by leaving the bicycle in this place, there is a signal that your debt and your use of your bike are finished. So, I think there is a QR where you put the camera of your mobile and at that moment your time on your use of

this bicycle is finished. […] This is the occupation of the public roads by public works. This is the first one. You have seen this one. On this website, you have here the weather, quick information about COVID, about the pandemic, and air quality. You have here the level, some standards in, for example, Avenida de Francia or Boulevard Sur. I do not know about chemics, maybe this is sulphur, NO2, ozone and there are some on particles (PM10). Okay. These are some data about unemployment in Valencia and employees. So, you have an employee feed, the people in December, in January, in February. So there is female unemployment, male unemployment and global unemployment. And you have here some tweets about traffic and other interesting information. Okay. So, this is a part of all the Big Data. This is the noise map.

Ella Pommeranz: Then what kind of information do you as a DMO collect? Cause this is from the city hall.

Joan-Carles Cambrils: That is from the city hall, but we do not collect this information. We just recommend in some places on our website that you have this information that maybe can be useful for you. So, one of your learnings today is that we have very good relations with the staff working there at the town hall with all these information and databases, but we have not integrated the databases. Not yet. We have a good proposal together. We have a good intention and there are common grounds about the importance of sharing this information and translating it into English, German and Russian, to different languages. But that is the level of the project for us.

Ella Pommeranz: But what about data from hotels? Do you collect data for that?

Joan-Carles Cambrils: Data from tourists, from tourism? I will show you another website for that. So on our website VisitValencia.com, not all the dashboards are in English, mostly in Spanish. Okay. But I give you a tour on this website and you can see what we have here. We have information about travelers and overnight states. This is the average stay of cruises. So, these three dashboards are linked to demand analysis. If you select one of these dashboards, as you can see, we are using Power BI. All of our tools are from Microsoft. […] I invite you to have a look at all these dashboards and I am just explaining to you how we collect this information. We collect all this information from national Spanish statistics. And these national Spanish statistics are delivered when they collect the data after the stays. After the information they collect from hotels, camping and other types of accommodations. So, the source here is a national statistic. We buy data for a local representation, for a locally acceptable sample. And they try to collect a minimum number of data to be statistically representative.

Ella Pommeranz: Okay. Did you say that you buy the data from the national statistics, or they provide them to you for free?

Joan-Carles Cambrils: We buy them because we want that they do a bigger sample because they just have prepared their statistics for being delivered at a national level and regional

level. Regional, or what we call in a Spain *comunidades autonomas*. That is the level of autonomy, like your states in Germany.

Cruises: we get information here from the port authorities. And we know by the Valencian port, which is the forecast for one year. The number of cruise ships that have planned to reach this port. But that information is not delivered here. Here, there is just information about the number of cruise ships that have arrived in the past. But we have the prediction for the whole year. You can watch that at the port of Valencia website. Here is some information about the past that we got from external sources.

About accommodation: the offer of accommodation, REVPAR or average price, this is some economical indicator, occupation, the rate of occupancy of rooms and beds. All these three dashboards are our feed by official sources. I mean the national Spanish statistics institution, *Instituto Nacional de Statistica*, is delivering this information, but unfortunately, it is delivered about one month to three months later. So, this is not Big Data at all. These are just some sources we have integrated into our information system, on just one website. And this is very interesting because many cities do not have the information so well organized as we have. I tell you an anecdote: an important hotel company from Barcelona called us about two months ago because they are looking for possibilities for investing in Valencia and these businessmen do not have a big staff searching for different cities and to understand how tourism is working in different cities, medium-sized cities in Europe. And when we send him to analyze just this data where you can have historical data on how tourism has improved during the last three, four years, except the pandemic period, but we have a good evolution from 2015 until 2019. So just with this data, they have a free and good analysis of information. And all of this information is on the same website. That it is quite useful for foreign investors, for local companies to understand how tourism is working. Not all of them are experts in statistics. Not a lot of them have analysts in their staff. They just want to follow some basic data about the city and tourism.

The online flats offer and online flats demand, all this information is bought for us from Transparent. Transparent is one of the companies that I think you mentioned in your examples. Transparent is a very interesting company. Maybe it is the leader in apartments, Airbnb, Hotels.com, Vrbo, all these marketplaces where you have these types of collaborative offers. Where you can rent your bedroom or sofa, all these things. So, we bought this data from Transparent, but we did a public bidding process and offer and they have lost the bid because of technical problems when they presented their offer. So unfortunately they will not deliver us this data again, but there is another company that I think is based in Canada, that is called Harmari. And they are a very good company too, but we have now some technical problems because they have to deliver their databases according to our structure of databases. According to our two ETLs [=Extract, Transform and Load], all the programming for bringing and selecting from the database the data as we want it to be displayed. So, we are working on this just in these days.

Ella Pommeranz: One question. When we talk about the VLCi dashboard, we are talking about this, or are we talking about the one from the city hall or the combination of all of these things?

Joan-Carles Cambrils: No, this is the tourism dashboard. When I talk about the dashboard, I am talking about this. I mean, this is not statistics. You can create your own reports and it is free. At the moment it is an offer free to everybody, even to you. Visit Valencia is legally a foundation and we have about 400 companies in the city that are members in our foundation. Our foundation is mostly supported and financed by the city hall. If you want to be a member of this company, you have to pay a fee, if you are a private hotel or a restaurant or a rent a bike company.

This is for example, the *apartamientos turisticos en Valencia*. And these are the apartments that rent a full house, not just rooms. You have a map here with the location of the apartments. Well, there are not so many apartments here compared to other cities in Northern Europe, maybe in Germany. I will tell you why. We have just about 4,400 apartments. The global amount of houses and apartments in Valencia city is about 400,000. So, it is just 1% of the global amount of housing and apartments in this city which has about 800,000 population. The average size of a family or residents in a home is two people. So, 800,000 population, 400,000 houses and 4,400 [touristic] apartments. In what marketplaces: as you can see Airbnb, Booking, HomeAway and TripAdvisor. They have been separated to avoid duplications by the Transparent company. So, this is completely Big Data analysis. But we do not create these dashboards, scrapping and preparing our own program for these marketplaces. It is an outsourced task. All this work is done by external companies, by the Transparent company. And the Transparent company sells us their analysis. We connect through APIs [=Application Programming Interface] or with ETLs to the database that they deliver us permanently. Here, this is the offer, but if you go to online flats demand, this information is changing continuously. More or less, Transparent has developed a scraper program that is reading permanently in these four marketplaces. Four or more, there might be maybe a small marketplace where they have access in different countries. Okay. But when they are scrapping and they identify an apartment here in front of my office that is offered, they know that it is free to offer in different moments and what is the price that they ask for renting the house. If this scrapping machine is permanently scrapping at high speed and sees that this apartment is not rented, one minute later, they take the information about which was the last price this apartment was offered. So, they will know at this moment, which is the rate of occupancy in the city, which apartments are free, and which apartments are occupied in which neighborhood of the city. And all that information is the information we buy to be delivered on our website. But it is not made by our own resources, it is not performed by our colleagues, by our officials, our workers. So, this is the most efficient way of having this data. You buy external sources, you buy the information that is delivered by specialized companies and these companies specialized in apartments, they do not have the ability to work with cruise ships or with airplanes' flight capacity or train capacity. And there are

different suppliers in the market that every week, every three weeks send us emails "Would you like to have data about my consumers? Would you like to have data about telephone companies and mobile users, et cetera?".

Ella Pommeranz: And what type of companies are these? What are their names, if you are able to share names?

Joan-Carles Cambrils: Well, for apartments companies, I think the best is Transparent [in terms of] price and quality, but this Canadian company called Harmari gave us a cheaper offer.

Ella Pommeranz: And what about MICE data or cruise tourism data?

Joan-Carles Cambrils: Eh, on MICE data I could not tell you the name of the companies now, but there are companies. If you browse maybe the International Congress and Conventions Association, ICCA, this is the most important network of professionals. Maybe there are suppliers of statistics and they have linked their databases to some venues all over Europe that can deliver you information about the type of visitors that are moving, but we do not have information about that.

So, this is, as you can see, this is exactly what we understand as Big Data. If we go to air traffic or regional airline capacity, air connections or tourists' profiles, all this information is delivered by another external company. These are predictive dashboards too because they are showing you the interest or the capacity offered by airlines for visiting Valencia from about 82 different airports that are flying with direct flights to Valencia and all this information is delivered by ForwardKeys. They are the best providers of airlines, traffic, capacity, prices, and they have much more information that we do not show. And we have some other further information, but maybe you do not know that ForwardKeys, which is a global company working in 70 countries, the headquarters are in Valencia. That is because we have a very good relationship with them. We have a win-win agreement that is not available with other cities. So, very special conditions, because they try with us and they have meetings with us to understand tourism companies' needs. Because the problem with Big Data is that there is a big amount of data that is completely useful for taking decisions or very expensive for taking decisions at a level of small- and medium-sized companies or small- or medium-sized destinations.

So, I recommend you to see which is the capacity from Cologne to Valencia. In the airlines market, there are two big seasons, what they call the winter season and the summer season, I mean, six months is one season. There is autumn and winter, and spring and summer is the summer season. You can have information on the present season or the coming season which they call Summer 22. But I am just talking about the present season market in Germany. Frankfurt is the market with 231 flights, so it is the market that has grown the most. We have 77 [additional] flights from and to Cologne. These are the winter flights, so consider autumn and winter. It has been planned until the Easter weekend or the end of April. 88 flights have

been planned for maybe 25 weeks, that means [there were] maybe three flights per week you could get from Valencia. And it has increased 700% compared to the previous season, but that is an easy explanation though. There were no flights [because of the pandemic]. There was one flight a week. And now you have three, four flights a week. And I can get further information. So I choose Cologne-Bonn airport, and the story is: 88 flights. As we know, it has increased by the number of 77 flights compared to the previous season. Here we have winter 2019, summer 20, winter 20, summer 21, and winter 21. These are flights and these are the seats. So, there are 16,000 offered. And we know that all of them are managed by Ryanair. […] So, this is a very interesting tool for us because we have information about 82 different incoming markets. Places where we can prepare strategic meetings with local travel agents or in those markets where we see that the company is not offering too many seats or flights and we do not want to lose that connection as a city. For example, with Riga or with Oslo or some other places where companies used to go to Alicante airport, south of Valencia, where they have beaches and they have all this sand and beach offer. But if we want to travel with some strategic connections in order to give some breadth to the company, we contact the DMO and we do some marketing actions in the city, or we publish information in local newspapers, or we invite some influencers to visit Valencia in order to improve and to protect that airline connection. Because airline connections are key elements for urban tourism. So, I gave you an itinerary over the city dashboards and tourism dashboards that we have not integrated at this moment.

Ella Pommeranz: So, basically you are collecting all of this data and you are buying the data from external providers and then you said that you are providing this data to tourism businesses for free. Is that correct?

Joan-Carles Cambrils: Yes, it is true. We are thinking in the future when we finish the, well we are now in the third phase of developing our tourist information system, we want to work with better information about the demand side of the market, about the visitors. And we are trying to get better information about flows, linked to a project funded by European funds, that is managed by the provincial government. The provincial government wants to set up 400 Wi-Fi sensors to give free Wi-Fi in different places of the province of Valencia to understand better the internal flows. And we would like to have better information from another demand study, another survey. It will be through a survey. We have to complement information with a traditional way of getting information. That is the survey. So, now we want to do the next survey. It is a big survey with the sample minimum being 2000 valid questionnaires [filled out] in different moments of the year in different places of Valencia. And it is an important and traditional source of information asking about demand, characteristics, consumer behavior, and level of satisfaction in different services, private and public services of Valencia. Characterization is: where do you come from? How did you get there, how many times have you come back to the city, which is the size of your group or family and consumer behavior. We used to do this survey every two years but because of the pandemic situation, we postponed it. But we would like to do another one this year,

considering that we did not expect the war. It is a big problem, but that is not important. How the market will [evolve], I do not know, but we will contract another local survey with plenty of information that we need, and we can get this information from other sources. Not with a good, statistically significant representation compared to the universe of visitors we have. We will pay for this and it will be prepared by an external company. But all this information will be delivered in the same way as you see in these dashboards. Okay. But that will be static information.

Ella Pommeranz: Then one more question. Do you also provide reports where you present and analyze the data in reports and send them to some businesses? Can you maybe tell me a bit more about that?

Joan-Carles Cambrils: We do not use reporting. Usually, we prepare a summary of different indicators and we deliver it to our members always every year. It is like a summary of a temporary analysis or the evolution from year to year. We could do reporting and get paid for hotels or restaurant chains or some other local companies, we could do that. We could even deliver the computer system design, to, for example, the city of Malaga. They would like to have our structure, our design. We could get a lot of money for that because we have invested in all that you have seen there, counting the cost of sources and the cost of development, more than 400,000 euros.

So what is the problem with Big Data? The problem is that the cost of acquiring, of buying information is very high for small companies, and even for cities like Valencia. I mean, we have money, but sometimes banks, mobile companies or other researchers that analyze social networks or blogs or websites for sentiment analyst or opinion analysis, the information they supply us is not complete. It is not relevant many times because there is a lot of work to do with the good interpretation of this data. I could give you a few examples because we have contracted Mabrian, that is a Spanish company and you have mentioned it. We have contracted them in 2019 and they gave us some information about sentiment, happiness, disagreement, and agreement. They used to get information at that moment from Instagram and Twitter. They could not get information from Facebook. Facebook is much more difficult. But when you analyze the results there is a problem with the IPs or with the communication company or where you have your cloud or where you have the servers. So, the sample shows us that there are a lot of people from the United States, and it is not according to the different types of tourists that we get from the national statistics. Or there are many people where they read their Twitter or their Instagram and it shows that they are from Valencia, but from Valencia in Venezuela, because in Venezuela, there is a city called Valencia that has 2 million population living there. 2 million, much more than Valencia in Spain. Or some other type of analysis that has to be cleaned before they use it in the reports. So sometimes they give you information, but it is not [accurate].

The problem with cities in Europe or cities like Valencia is that the cost of sourcing data is quite expensive and companies do not know exactly which is the value of that information.

That is typical in banks, banking information, paying cards like Visa, MasterCard, or mobiles information. So mobiles banking information, visa information and traffic information, some people can get information from mobiles in the open areas. All of this information is quite expensive. I have been in many meetings where they come to you […] and they will deliver you information about consumer behavior of visitors that have or use a data phone that belonged to this bank. Okay, so they can have information about the users of banks, of foreign banks from Deutsche Bank. Maybe a user of the dataphone that they have in the shops in Valencia, the Spanish bank has delivered a reader, paying card reader. So, if a person from Cologne using a Deutsche Bank card is using this device that belongs to BBVA or Santander bank in Spain, they can get this transaction and analyze it. But this is maybe 10% of the city, 8%, 9%, 11%. So, which is your share in Valencia? If there was a pool of banks integrating all of the information of foreign transactions in this type of goods, in this area of the city and so on, we could have better information, but we do not have that now.

So, I feel that where we need improvement, where we need a better quality of data in the future is consumer expenditure, consumer behavior. We need a better understanding of sentiment analysis. Many of the suppliers in this field have scrapping programs giving you, for example, rankings: how do they rank by stars, what rank is better - scrapping in TripAdvisor or Hotels.com, in Expedia and so on. But I am not completely interested in the ranking from one to five or one to seven stars, because it has been empirically demonstrated by research that when people are happy, they usually do not take time to rank. When you really dislike completely a service you have received, you want to revenge yourself. So is a natural behavior of human beings. This is well known by independent researchers. We need to improve in understanding better what we call semantic language or opinion analysis. If this person put comments and put written evaluations, what is happening there. The average of the evaluation or the assessment of visitors is not so important. It is much more important to have scrapping programs that read this conversation and give the right interpretation of what is happening. Maybe a person that gives one point from seven to a destination, maybe it is just because he had a bad discussion with another guy or with a guardian or with a steward there. It was a discussion, but you are not telling me [anything] about the quality of the museum or the park or the experience. It has been just a small incident that made these people so angry.

Ella Pommeranz: It is kind of funny because I used to work in quality management for a hotel in Spain. So, I was also looking at like online reviews and also I did feedback talks with the guests. And when I looked at the surveys at the end, I was not really looking at the ranks from like one to five. I was also just looking at the comments and then putting that into categories and analyzing that. So basically, just what you said.

Joan-Carles Cambrils: Yeah, because it gives you much more rich information than being ranked eight or nine on Booking.com. But for Big Data, we need a tool or we need a solution that can read thousands of comments and can [provide] an accurate analysis of data. So there are a lot of researchers working in what they call natural language.

Ella Pommeranz: And do you think Valencia will internalize this process or do you think they are going to outsource it to another company to do that?

Joan-Carles Cambrils: I can tell you that 90% of cities in Europe like Valencia can not pay for a big internal analysis team or a big statistical team. We have three, four people, no more than that are working. We have a statistic person on our staff. We have a junior project manager. We have a computer systems and computer engineer working, giving a hand in understanding the programming or the right connection of databases or ATLs to dashboards with the infrastructure. And we have a coordinator of intelligence and data that is Ricardo Millet, but we have to be agile. We have to be more efficient than just having interviews with many suppliers, understanding what they are offering us, evaluating to see if these suppliers are serious people because they will supply you with data in the future. I think in the coming years, there will be specialized companies like ForwardKeys, Transparent or Harmari, or some others in specialized fields where each of them has improved the research, the accuracy of data, the cleaning of noise and what is not relevant. And when they prepare a good exposition, a good example of your city you will buy them. But they should come here. What we have learned is that there are a lot of people trying to sell pilots projects. No, because scrapping and reading social networks, I am sure there are even free programs that will do that for you. Programs that read websites and give you information about that. That is a free resource you can get on the internet. But the question is to discriminate between what is noise, what is an accident and what is essential and what do you want to read?

So the future of the big companies or phone telephone companies, bank companies, transport companies is well organized. But we need information about consumer behavior in the city, setting up thousands of sensors or combining data collected from mobiles and public scans that can read your IP indicator or the number of your phone, if you have the GPS open or not. They can just read some information about the mobile, not much information. But we need information about flows, about consumer expenditure, about sentiment analysis, about traffic jams and mobs, where many people are together and going together in different places. That is all we think, realistically, we could have in the coming four or five years. We need to improve a lot the connection. In many cities, we have the tourist information data on one website and the local data on another one. It should not be expensive at all to exchange with APs and give information from one to the other and use the local information for tourists. We are going to work in the coming years with that because they [tourists] start asking for us much more information about environmental indicators. That is much more important now with climate change and the problems we have with weather and pollution, and so on. And finally, the cities in Europe need to improve our services we have, what we call scalable projects. The German national authority or the Valencia regional authority should create these dashboards. They should buy the sources of information. And all these dashboards are completely well adapted to manage local information, regional information, provincial information, state information, and probably national information. And there are a few projects coming from the Europe Union. From EUROSTAT. I have heard about

projects of the Spanish national government. They have to allocate a hundred million euros just this year in a Big Data office where they want to buy information on, for example, apartments. It is the same cost for the Spanish authority to design a platform that gets information from those networks or to buy information from Transparent and displaying this information at a local level, at a regional level or a national level. Because the scrappers, the programs and the analysts in Transparent, they give you the data at a local level, at the street level. So why should we pay for this information from Valencia, from Transparent, if there is a big buyer or a big developer of their own programs. And when you browse and you scrap all the information on the database, that has well-built ETL, it is brought to the Valencia level, Alicante level, Madrid level, and it is the same cost for everybody, the same platform. The problem is that our regional government has developed not many dashboards. It is still developing this way. Benidorm has other dashboards, but we are the most advanced city just with these small and clear dashboards.

And what we can learn from Valencia is "please listen to your local companies". Your local companies do not need too much data. They need to understand how many tourists will come in the coming years. So, in the coming months, when we have the information from ForwardKeys, we will get more information about the train stations. We have the cruise ship information, but it is not in the dashboard, it is a preview, but we have the capacity, and we will have the record of cruise passengers this year, according to the offer capacity. I do not know if people will buy tickets or will do the holiday trip on the cruise ship. I do not know, but we know the offer [=supply].

So, people want to know how many people will come, from where they come from. How much money do they have and how do they spend money in the city? So you do not need thousands of data, because you do not have time every day to analyze it. I mean, my time is limited. But it is because I and the local companies only have a couple of hours a week to check how are things, to have an impression, to have a discussion among your colleagues in the other companies and that is all.

Transcript 4 – Written interview with Francisco José Garcia Martinez (Head of Strategic Planning at Turisme Communitat Valenciana) on behalf of Mario Villar Garcia (Head of Touristic Intelligence at INVATTUR)

Date: April 7th, 2022

1. Data availability and access

- What type of local/national data (= information generated and collected in Valencia) does your organization have access to and use in its operations? (e.g. overnights, arrivals, purchasing behaviour, movement patterns of tourists)
- Do you know the VLCi platform or Valencia Al Minut dashboard and how are you using them? I've installed and test for 5 months, there're two options, for citizens and tourist. But the information is the same, so the problem is first the confusion with that two different roles, and second, the same public image for making administrative task and smart information.
    - Are there any functions that are missing? Information relative to special promotion offers to tourists, Information about security travel related to the specific area where the tourist are. It´s a shame because Valencia Council has a great geographic system.
- What type of global data does your company have access to and use in its operations? (e.g. Mabrian, STR, UNWTO, AllKeys, etc.) Mabrian, UNWTO, google trends, google insights.
- Are you co-sharing your data/ dashboard with other organizations or companies at the destination level? Why (not)? Not on-line, because administrative problems with our Big Data projects, but we share our information specifically to all administration levels.
- What type of data will be necessary to have access to in the future? We need information in two directions; first direction, in collaboration with smart destinations like Valencia or Benidorm, data from the local on-line sensors which produce real time information about use of resources, security, hospitality opportunities; second, we need global information about the mobility, payments and use of digital resources, like a urban mobility card. The integration of all this amount of data it will be serve to the tourist integrated with a net of beacons to promote tourism as a hospitality industry, what is the point of interest which converts Big Data on valuable information.

2. Data analysis and processing

- How is Big Data used in your organization?
- What type of basic analysis are you using to analyze and learn from past events in order to be reactive? (e.g. reporting, KPIs, dashboards, scorecards) reporting, KPIs, dashboards
- What type of advanced analysis are you using to analyze the current situation and to predict future events and trends in order to be proactive? (e.g. data mining, simulations, predictive modeling) predictive modeling about the interest of tourist researches related to the local weather, and the correlation from airport capacity and researches related the weeks of the year.

3. Infrastructure

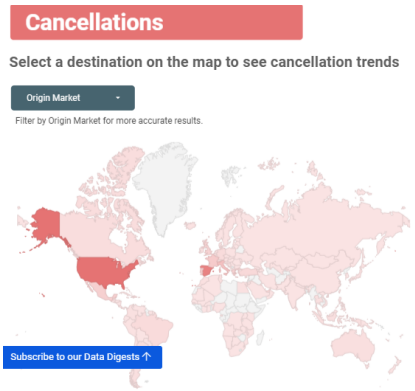
- What type of infrastructure does your organization currently have that enables the analysis of Big Data? (e.g., cloud storage, internet connection, sensors) N/A
- What infrastructure would be needed for an optimal use and analysis of Big Data (now and in the future)? cloud storage, cloud computing, sensors/beacons

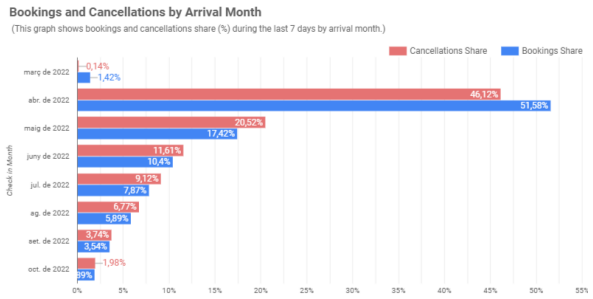
4. Employees' skills

- How would you describe the skills that employees in your organization currently have in regard to Big Data analysis? Even though this techs are generalized in our days, we don`t have included in the organizational culture as we want to introduce the hospitality as the most important value in smart tourism.
- What other skills would be needed for an optimal use and analysis of Big Data (now and in the future)? Practice makes the master, but we don`t have time in our ordinary days, so the key it's migrate our day-by-day operations related to techs which offers solutions by analytics, for example convert actual administrative systems to block chain contracts so the result of operations are directly useful for Big Data analytics.

5. Opportunities and challenges

- What are the main opportunities of Big Data for the tourism sector? To promote tourism as an hospitality industry
- How can Big Data help to cope with different crisis coming from the broader environment (safety and security, health, economic, demographic, climate etc.)? The amount of open information like Our World in Data, it`s very important to make a quick response, for example with Travel Industry COVID-19 Monitoring (travelgatex.com) you can control the national % of cancelations for the most relevant markets, this variable it´s now relevant in our days, so it´s a test to measure the impact of new crisis.





- What are the main challenges related to Big Data implementation in the tourism industry?

6. Future outlook

- What are some key factors that destinations or businesses need to optimize the implementation and performance of Big Data? The cloud computing and the quick developments, without tools and adequate engineers and data analyst, we'll arrive late with no return of value.
- How do you think will the adoption of Big Data in the tourism sector evolve in the future? it will pass from clients analytics to organizations

---

Transcript 5 – Interview with Michelle Wols (Head of Marketing at ForwardKeys)

Date: April 8th, 2022

Duration: 38 min

Ella Pommeranz: Okay, so, I thought before we start it would be great if you could give a brief introduction about ForwardKeys.

Michelle Wols: Okay, so in ForwardKeys I am the Head of Marketing which means that I do most of the product descriptions, but also visit a lot of events, and talk to the clients to learn more about their needs. And what FordwardKeys does, it is a data analytics company. We started off as a purely aviation-focused analytics company, which means that we have data on all of the global scheduled flights. And we also have about 60% of all of the flight bookings made worldwide. And now we also have flight searches. That means that with the information we have, it is all anonymized so we do not have any personal data, we can detect trends. We know which nationalities book how long in advance, we know where they go, we know how long they stay, we know if they visit multiple destinations in the same trip or if it is only one. And that is the kind of information that we primarily sell to destinations. We now have about 65 to 70 destinations as clients. You can think of destinations from the Japan national tourism board to the Americas and everything in between. It is a pretty big market for us. The biggest actually that we have. And what they would do with the data right now is mostly marketing. So, decide which audiences to focus on. If they know that there are already flights between their destination and another, they know that that is there is where they have to do a marketing campaign. And with our data they can see when they should run those campaigns and what periods and obviously the destinations' objective currently is to keep tourists as long as possible and to attract more tourists that have a big economic impact

but leave a very low social impact. And well, that is what ForwardKeys used to do and now we are also moving into more of this phase because we see that there are a lot of destinations with overtourism. Think about Barcelona and Amsterdam, for example. These destinations are no longer looking into attracting as many tourists again but they are trying to limit those tourists that cause a lot of harm and attract more wealthy tourists. And so we are also looking into other data sets that would help us understand where in the destination do people go. So that in the destination they can try and spread out those tourists better. And also monitor if their strategies are successful.

Ella Pommeranz: Okay, great that sounds very interesting I do have a lot of questions regarding that, so I am going to get back to some of the things you already said later. And if you do not mind I would start off with the interview questions that I sent you. So yeah first, I wanted to talk generally about the opportunities and challenges of Big Data. So, in your opinion, what are the main opportunities of Big Data for the tourism sector?

Michelle Wols: I mean data is very new and what we see is that right now a lot of the destinations use it, because a lot of them are publicly funded with government money, so they have to report on how money is spent. And so, what we see happening a lot is that most tourism boards would actually only use our data to report. So that would be "Okay we have seen that this happened" and they use the data to explain that that happened. And I think a very big challenge for tourism boards, or for any organized organization globally, actually is to use predictive data. So, to use data to predict the future and change your actions, depending on what you see. I think that is probably a very big challenge. And then, tourism boards have so many challenges that they are facing in which data could actually be of help. Think about sustainability. I mean long term it is to question if we can still keep flying around the world, in the same way that we do now. And so, if we are going to limit that which is very likely going to happen, the tourism board still has to decide which tourists they are going to focus on. Yeah, I think that would be one of the biggest challenges that the tourism boards face.

Ella Pommeranz: Great and you already mentioned that by looking at this data, you can also see what type of tourists are coming so then the destinations can plan ahead like "Okay, we need more people working at this time, we need to offer our information in these languages". I think this is also some of the opportunities the data can maybe help with.

Michelle Wols: Yeah, and if you have looked at Visit Valencia, this is probably one of the big tasks that they have for themselves, but that depends on the size of the tourism board. So the national tourism board would not do this but the city DMOs, like Visit Valencia, create dashboards for businesses in the destination to help them make use of it as well. And I think that is, of course, the second challenge if the tourism boards themselves find it sometimes hard to make decisions with data, then how can you get the local businesses to do so.

Ella Pommeranz: I think a lot of the tourism businesses do not know what to do with the data that they already have, right. So I guess someone needs to educate them like there needs to be more training, but we are going to get to that later. Then I would move on to the next question which is about crises. So, I think we all know by now that crises can hit us anytime. We have seen it with the pandemic, we have seen it with the war, it can be anything. So how do you think Big Data can help us to cope with these different crises, so that could be like safety and security, health crisis, economic, and so on.

Michelle Wols: I mean what we saw is that COVID was for ForwardKeys as a company both a curse and a blessing because obviously there were no travelers out of the blue. And we thought we would be losing a lot of clients because, why would you need to buy data to know that no one is traveling. Because borders were closed, no airplanes were taking off for quite some time. But actually, we won a lot of clients, because all of a sudden, there was this need that people needed to know what was actually happening and if tourism would start again. And we see that actually, the war in Ukraine has had exactly the same impact. We get a lot of people reaching out saying "What is happening? Are people because of the war in Ukraine not traveling to Europe anymore?". Because that is the thing that we have seen in the past, for example when Ebola happened in Africa. I believe it was only really a few countries where Ebola was really present but people stopped traveling to the entire continent. And so now, the question was do we see that people stopped traveling to Europe as well, because of the war. Because it might seem very close, you know, you do not travel to Spain, because in Ukraine there is a war. And we did not see that this time so data can really be of help in this, in understanding how are consumers responding. Do they still want to travel, do they not? If they do want to travel, then, have their preferences change because, in the end, businesses have to adjust.

Ella Pommeranz: Very good, and as you already said, there are some main challenges regarding the Big Data implementation in the tourism industry. Can you maybe elaborate on that further?

Michelle Wols: Well, we have already covered a few already. One that I would add is, it is sometimes called data swamp, where you have loads and loads of data. You have many different kinds, I mean we are only one of the providers to destinations. They usually also buy hotel occupancy, for example, or Airbnb bookings or geolocation-based data, social media data. They all serve different purposes, but also sometimes are quite aligned. And, in some cases, one data set can show a very different picture than another because of coverage issues or because it covers different kinds of travelers. I mean getting an understanding of which data set to use in which context, and how to read this information I think that is going to be a big challenge managing it all.

Ella Pommeranz: Is ForwardKeys doing something in that regard, like educating the companies you work with, or is that another competence?

Michelle Wols: We definitely try to help our clients do more of this but that is a slow process and also quite often, they have to change internally. Because, in some cases, they only have one data scientist, for example. and at FordwardKeys, we already have more than 20, I believe. So I mean if you really want a company to be data-driven that just means a lot of investment.

Ella Pommeranz: Okay I am going to go back to the data that you are using and have access to. We already talked about that you are using air travel databases, so you have data from travel agencies and airlines, correct?

Michelle Wols: Yes.

Ella Pommeranz: Okay. So, what type of data does your organization have access to and use in its operations?

Michelle Wols: I mean, in our case, we have the three data sets that I mentioned, so that would be the tickets booked, flight schedules and flight searches. Those three we currently sell. What we do with them is quite broad actually. We hardly ever sell raw data. We like to make it as easy as possible for clients to use it, so that means that we combine data sets as well. To give you a practical example, one of the tasks that destinations also have is negotiating routes with airlines. And so, they need to understand what routes currently exist, which flight schedules, or what flight schedules currently look like to my destination, but also my competitors. And then we combine this with the searches, to see if a route is underperforming. Let us say that a route can be empty every week. We do that by combining the schedules with our tickets. And we make an estimation, as to how full the airplanes actually are. If a route is underperforming that gives clues to the destination to actually talk to the airline or to launch an additional marketing campaign to make sure that the airplanes are full. And when we combine schedules with searches, we can see if there is, for example, demand to fly from Amsterdam to Valencia, but there are no actual flights. And so we combine that for the destinations as well. And the way that they can consume this would be super broad: we have platforms, where they can access the data, already provided in easy-to-read graphs and tables. But they can also get the raw data and then do it themselves.

Ella Pommeranz: So you are selling both the insights and then the data as well?

Michelle Wols: Exactly

Ella Pommeranz: Okay great. And then are you co-sharing this data with other travel intelligence platforms, like, for example, Mabrian, STR or UNWTO?

Michelle Wols: UNWTO is one of our partners so they actually have their online dashboard for free, some of our data is in there. Mabrian is definitely a competitor, they do have some flight booking data as well, and I believe they have flight schedules too. But then we are very different companies. Like ForwardKeys strives to have very high-quality data, we have

very high coverage of all the tickets sold globally. Mabrian would not have that. But they do provide other data like spending data from MasterCard, they also have the hotel occupancy and they actually started off as a social listening company, so it means that they do a lot with social media data.

Ella Pommeranz: Okay, good to know. What is your relationship with Visit Valencia? So I know you are working with them, right? And you are also headquartered in Valencia?

Michelle Wols: Right exactly. Valencia helps us in product development, they are participating in every single study that we do about developing our products. I mean it is very easy because of the geographical proximity. And we have an especially good relationship with a guy called Jaume. I think he was the one to accept the price of Valencia as an innovative destination. And what he focuses on now is one of those challenges that we mentioned before, so it would be about measuring how sustainable the destination is. And he listed, I think, 45 KPIs, ranging from water quality to carbon emissions to anything that you can think of. And so, they are helping us, because we would like to enter more into that space as well. So, they help us understand what exactly does the destination already have in terms of data. And what can ForwardKeys do to make it easy to measure between destinations how they are doing in an objective way. Other than all of the destinations doing this themselves.

Ella Pommeranz: That is great that it is like a reciprocal relationship, so it is like giving and taking. Okay, then the next question would be what type of data do you think will be necessary to have access to in the future? Is there something that is missing that you think will be important in the future?

Michelle Wols: I think there are two very important points in destination marketing organizations that currently do not exist. I do not think there is new data that has to be used for it. I think it is all there, it just needs someone who puts it all together. I have talked to you about the economic, social and environmental impacts of tourists in the destination. And especially there are two points, like on the economic side of things destinations want to make sure that as many people in the destination win from tourism. So, destinations want less of the people that cause a lot of trouble. We always use the example of the drunk English visitors. And then they all want more visitors that have like a very low impact and that actually do not go to all of the tourist hotspots but spread out over the city. Right now, there is no uniform way for destinations to measure these kinds of profiles. And to actually use that kind of information in their marketing campaigns. I mean there are some trials, but there is not one system. And then the second one would be measuring the sustainability of tourists, so what is their environmental impact. How many carbon emissions do they emit to do to travel, and how much water do they use. That is something that just is not there yet.

Ella Pommeranz: And do you think ForwardKeys can help with that or it is like a separate issue because it is not really related to airline insights?

Michelle Wols: I mean we are trying to develop a way to do this. But we are far from having accomplished anything. It will be the earliest next year that we will actually be able to show something.

Ella Pommeranz: I mean it is a process, it is not that easy to just implement it, I guess.

Michelle Wols: Yeah, exactly because also it is one thing if we could develop this as a holy grail and give destinations a uniform way of benchmarking themselves with each other. Uniform but also objective, I think that is also important. The first step is developing, and the second is convincing destinations that we have developed this Holy Grail.

Ella Pommeranz: Okay, great, then I would like to move into how you use all the data that you collected. So, how would you say in general is Big Data used in your organization?

Michelle Wols: Well, in our organization it is obviously used quite a lot. We analyze and use our data, but only for the benefit of our clients. So it is like a service that we sell. If I would say how we use our own data internally, then we obviously as a marketing team, we also buy data on how often we are mentioned in the press and how many people visit our website. We have data from our marketing campaigns. The sales team does exactly the same, well we do it together actually. As a data analytics company, we obviously also have to be proactive in this, we cannot tell our clients to be using data and then not do it ourselves.

Ella Pommeranz: That is true. Very good point. Okay, and what type of basic analysis are using? I saw that you offer dashboards, travel statistics as you already mentioned and then also APIs. So, what other kind of basic analysis are you doing at ForwardKeys?

Michelle Wols: We, of course, have a lot of data scientists for internal research before anything is implemented. So we do lots of correlation analysis and we build models and it is a lot of time that we spend on that. And then for our clients, we either sell flat files, so that would be just data files they can buy and then we also indeed have different kinds of dashboards that they could get access to, or they can even connect to our data via API. And that would actually be about it for our clients' options. Well, I mean very rarely we also handwrite reports. For those destinations that are beginners at data analytics, they do not have a lot of experience. So we really need to take them by the hand and in this case, someone from our insights team would write a report. Or they would do a presentation in which they would explain the trends to your destination look like this or this. And then brainstorm with the destinations to see "How can I use this information?".

Ella Pommeranz: It sounds, again, more like selling the insights rather than the data itself, right?

Michelle Wols: Exactly yes. Absolutely.

Ella Pommeranz: And then what type of advanced analysis are you using? So, correlation analysis is already advanced analysis, and then I saw on your website that you offer BI to predict demand, right?

Michelle Wols: Yeah, so the two I mentioned. We do a lot of correlation analysis but that is mostly because we try to correlate our data to others. We also work in a few other markets, for example, travel retail where it is all about the tourists that buy something at the airport. And what we try to do there is to connect our data to the sales data of the clients. To see if we are able to predict their future sales, which we can. But that is something that we would do depending on the client. We do have models and forecasting models, those are probably the most advanced stuff that we currently do. So, we have one model to predict or to estimate the entire travel market globally. Like I said, 60% of bookings [worldwide] are tracked by us but then that 40% of the [rest of the] market, we have no information on. So we have a total air market model that estimates, with a lot of detail, how many people are traveling in the world. That is mostly because right now low-cost carriers, for example, are impossible to track, they do not share their data with anyone.

Ella Pommeranz: So, usually you buy the data from the airlines directly and from the booking systems?

Michelle Wols: Almost. We buy our data from IATA, and so IATA gets all of the data from the travel agencies. And they have a project with airlines where airlines share their data, and then IATA also shares that with us. But Ryanair and AirAsia, for example, would never participate in anything like this. So, what we do with this model is to estimate on a daily basis how many people traveled with Ryanair in Europe and what kind of people is it. So all of the details that I mentioned before, like the length of stay, how long in advance people book, that would be things that we estimate. And then built on top of this model, we also make forecasts where we estimate how many people will be traveling three months, six months or a year from now in the future. Three months in advance, we can predict with quite a lot of detail but obviously a year in the future we do not because…

Ella Pommeranz: You never know what is going to happen.

Michelle Wols: Yeah exactly.

Ella Pommeranz: Okay. Let us move on to the potential barriers of Big Data analysis. And one of the ones that I found very often when I was doing my research is actually infrastructure. So this could be the cloud storage, internet connection and everything you need for doing Big Data or just data analysis. So what type of infrastructure does ForwardKeys have that enables the analysis of Big Data?

Michelle Wols: Well, we have both our own servers, but we also work with the servers of Amazon. I am talking about thousands and thousands and thousands of bookings every day. Now with COVID it has been less, but after COVID it will be the same. We are talking about

a lot of data that needs to be stored. So, we obviously need a lot of capacity for that. But our clients, unless they buy the flat files, they do not have to have it themselves. And even if they would buy the flat files, it would still be a limited amount of data that they would actually have to store themselves, so the issue would be less big.

Ella Pommeranz: And then is there anything that you need that you do not have right now, but would be optimal to have for analysis of Big Data, like for now and in the future?

Michelle Wols: I do not think I am the right person to ask for this. I think we have everything. We always have the newest technologies in terms of data, but I do not know if there is something really missing in the market or anything.

Ella Pommeranz: Do you think that maybe destinations will build up this infrastructure in the future, or do you think they are going to continue to outsource these kinds of things to external companies like ForwardKeys?

Michelle Wols: I think it will be outsourced, to be honest. I think it would be a lot cheaper for them to outsource it because we as a company are very prepared and they are not. I think only with the human capital, it would already be a big investment on their side to be able to use data. I do think it is something that they will outsource, that would make more sense.

Ella Pommeranz: I mean it makes sense, like you already mentioned that you have 20 data scientists alone and most companies only have one. I do see the point. If we are already talking about human capital, how would you describe the skills that the employees in ForwardKeys have or in the tourism industry in general?

Michelle Wols: Well, I think there is a big difference between ForwardKeys and the industry in general. I would say that in our company the level of education is really high, for obvious reasons. In travel itself, it really depends on the organization. I see a lot of destinations that have started to hire PhD students from universities to get to the same level which I think is a good thing. But I mean there is also countries where, for example in Spain, every single civil servant gets a job for life, so there are a lot of people who have been working in the Ministry of Tourism or in destinations that have been there forever. And when they started working, the requirements were not what they are now and so there is a bigger challenge there and getting these kinds of people also educated and willing and capable of working with data.

Ella Pommeranz: Yeah, I think it is more about keeping up with the trends and developments.

Michelle Wols: Yeah exactly.

Ella Pommeranz: And what do you think what skills would be needed in the future in the tourism industry to analyze data?

Michelle Wols: What is missing?

Ella Pommeranz: Well, basically, just what you said, like basic data analysis knowledge.

Michelle Wols: Yeah exactly, I would say that that is probably the biggest challenge because I mean I did not graduate not too long ago like I think I graduated six years ago. And I had some analytics courses that you needed and even there people hated analytics. So yeah, that is going to be a challenge, like people have to realize that it is not that hard to work with data and use it.

Ella Pommeranz: Exactly. […] We are almost coming to an end, so lastly I would like to talk with you about the future of Big Data. So what key factors do businesses or destinations need to optimize the implementation and utilization of Big Data?

Michelle Wols: I mean there are a lot of things to be honest, like, I think, what is most important, as a start is a board of directors, that is willing and sees the benefits. If they do not support anything of this kind and if it is people who believe that with their gut feeling you can solve everything then nothing is going to change within the organization. No matter if there are people lower in the hierarchy that do believe that it is necessary. And after that, it is all about training. I think a lot about training. But first, there should be that willingness otherwise you are not going to get anywhere.

Ella Pommeranz: And then, last but not least, how do you think will the adoption of Big Data in the tourism sector evolve in the future?

Michelle Wols: I think it will be very similar to any other industry. Like I said, what we see until now is that a lot of people use data to explain what has happened in the past. That is the first step in starting to use data analytics. And I think once people get more acquainted with how to use the data, then, slowly but surely they will start to move into more of predictive analytics where they will actually look ahead and see how [they] can use the historical data to predict the future or to understand or anticipate the future. That is a slow and probably also very painful process where managers and directors have an important role to push those questions from changing within the organization.

Ella Pommeranz: Okay, great. Well, thank you very much, that was it already.

Transcript 6 – Interview with Jorge Ferrándiz (Technician in Sustainability, Innovation and Digitization Projects at HOSBEC)

Date: April 20th, 2022

Duration: 57 min

Ella Pommeranz: First, let us start off by talking about the opportunities and challenges. So, in your opinion, what are the main opportunities of Big Data for the tourism sector?

Jorge Ferrándiz: Nowadays, the main opportunity related to our sector […] is that it is making it possible to find behaviour patterns or consumption trends by integrating different data from different sources. For example, hotels now know… I am going to talk mainly about hotels because we are… Do you know HOSBEC as an organization?

Ella Pommeranz: Yeah, you are collecting data about the occupancy of hotels in the Valencian Community, right?

Jorge Ferrándiz: Yeah, we do a lot of things. I am going to share my screen. Because before starting the interview, I will show you some of our basic information or HOSBEC as an organization. So you know where you can find it in case you need it for your interview or as a profile of us. If you go to Hosbec.com it is our website, and it is like a really easy language, so the Google translate to English is fine. And here you can say first a quick history of what is HOSBEC which started as a hotel association more than 40 years ago and has been expanding with the years both in functions and in territory. Because we started only in the city of Benidorm but we are now in all of the Valencian Community with the three provinces. And also, we have here our services and you are going to see that we do a lot of things. I am only in the part of Big Data, but [we at HOSBEC do] a lot of things like communication, promotion, help services, we have a job bank. We try to provide all the services of the tourism industry that are needed in the Valencian community. And, mainly because of all of the services that we offer, it has been the main reason for us to expand in a matter of years because also the public sector has been seeing that we were doing really good work in these terms, and they promoted us to expand also to Valencia and Castellón. And also here, you can see all of our members of our association. And all of our service partners, and we have from revenue programs to formation like schools, tax services, everything.

Ella Pommeranz: How many hotels are you working with?

Jorge Ferrándiz: We must have like 200 or 210 [hotels] in all of the Valencian Community. But [counting] all of the businesses, because we also accept restaurants, apartment blocks. We expanded from hotels association to tourism association. And we have almost 300. I think we wanted to reach 300 now in May, now that we have our annual assembly. We are close to that. And here we also have the public statistics, which are mainly hotel occupancy. This is like the public side of our Big Data department. Then we have on the private [side] only for members or interest persons. But, for example, if you enter the hotel occupancy of Benidorm you can see the hotel occupancy of the city of Benidorm since 2015 by each half month.

Ella Pommeranz: Oh wow. And anyone can access it, right? Like it is publicly available.

Jorge Ferrándiz: Yes. You can even change the selection to compare, for example, the year 2021 with 2019 and 2015. Then we have all the other services [like] health services but I do not think that this is of interest to your research. So, this was a little bit about the [organization]. And what are the main opportunities that we are facing? First, finding the demand patterns in order to make the best decisions from an operative level. Related to this, the activity in the decision making, which is mainly supported by visual models. We have the data from before, like in 2014, 2013 we already had the data. We have been collecting the occupancy rate, for example, since 2006 or 2007, but it was really hard to understand the data. We like to say that we passed from Big Data to smart data because we now make all of that database that we had comprehensible and we make intuitive reports so that our members which do not always have a good formation, especially in this subject which is relatively new, they do not always have a really high formation about this. So, making it accessible for them and easy to understand for them has been key in the last few years. […] Big Data is like the whole, all the databases. But the data itself does not give anything; you need to transform it to get your conclusion […]. And then the process of digitalization is one of the main opportunities that Big Data gives us because knowing how the demand is going to behave, we can maybe reduce some of the processes which are not essential. For example, in a reception of a hotel, if we know that next week or next year, we are going to have double the tourists in our hotel, we may need to double our employees too. So being able to anticipate the situations is key in the sector. And also customer attraction and loyalty programs. Because having all of the information now, instead of occupancy, for example, having all the Big Data of customers, maybe in a CRM, allows businesses to make unique offers or to make unique services for one segment of customers.

Ella Pommeranz: So, the next [question] is about crises. I think as we all know by now crises, like the COVID-19 pandemic, for example, can hit us at any time. And how do you think can Big Data help us to cope with the different crises coming from the broader environment?

Jorge Ferrándiz: Big Data can help us in many ways. But focusing for example on the pandemic, as you [mentioned], we really use Big Data here in HOSBEC because we compare the occupancy rates that we were having in 2020, 2021 with the previous [numbers] pre-pandemic and that allows, especially the public sector, to give the adequate grants and subsidies to the companies. For example, we were able to know that on the books, we had for 2020 in summer, we were able to know in advance and knowing that the occupancy rate was basically zero because there was no tourism in 2020. That also gave security to the public sector in order to say "Okay we are going to give a grant for the next six months, or the next year to help our local businesses." Because it was the key objective of the pandemic that our businesses have to survive and a hotel has a lot of fixed costs. [So if] we do not have the support of the public sector, maybe most of them would have closed permanently because some of them close temporarily but without the public sector, probably, they would have closed permanently. And that was like one thing that helped us in the pandemic. But also in

the other crises that you mentioned, for example, the price of climate action, Big Data as a collection of data from multiple sources is a must if you want to know, for example, the carbon emissions. Also, you need to collect data from all of your different suppliers, your own emissions, you need to gather all of that, and make, for example, a carbon footprint report. And we are also [t]aking actions in that sense. We also have a members program where our associates can upload all of their information, it is just like an online form or an online survey that they have to complete and with that information and the electricity or gas fees and all that information we can [calculate] the carbon footprint. Apart from the carbon footprint, we also have like a custom certificate that says if they have sustainable practices in the business or not. That is maybe not so much [made] of numbers but we take Big Data in the sense of we are collecting data from a lot of different sources and gathering them all in one platform. Related to the challenges, one of the main challenges we have related to Big Data but also in general is the difficulty to attract new talent. The tourism industry, especially here in Spain, but I think it is a global problem, is having a lot of problems updating our employees. Like when a person gets 65 years old, and they have to retire, it is really difficult to find a young person that fits in the same spot, because education has changed, mentality has changed and not so many people now want to work, I do not know, cleaning houses. But it is a job that will always be needed. And this is something that access to Big Data does, especially because it is something relatively new. They are finding problems getting more students and employees that can fit in that role, so the businesses are paying for simple masters in Big Data or they are paying for like secondary education [for] their employees so that they are capable of doing these tasks. But also, apart from the talent attraction, building trust both with customers and suppliers, when you want to integrate information of a lot of different suppliers or even on your own customers in a system you always need the approval and, sometimes, especially with customers [you have to] explain to them why you need the data and for what you are going to use it. It is not always an easy task. So that is also one of the challenges: building trust. Capturing useful data because, especially in the last years or when a business starts with Big Data they think they need to gather everything again, and then will pick what is useful. But a good qualification also will save you time and money [in terms of] what data will be needed, and what data is useful for my reporting or whatever analysis I want to do. And finally, one of the challenges that we are facing is the changes in processes inside companies. For example, there is a big resistance to change in some areas of companies, especially when the employee does not understand the benefits of Big Data. Because maybe the top level of the company understands it, but if that data culture is not given to all of the members of the company onto the low members of the company, some of them will not be willing to do changes in their own work process, and therefore they will not gather the data needed or they are not going to do the steps needed in order to begin the process.

Ella Pommeranz: Who would you say may be responsible for initiating this change in the mindset? Do you think the companies should educate their employees or do you think maybe

there should be a government initiative starting education about Big Data and training? How do you think we can tackle that?

Jorge Ferrándiz: I think it should be more about like the own companies giving education to the employees because not every company will need to use Big Data. [It is] the same as now everyone is talking about blockchain, for example, but not every company will end up using it, [it is] the same with Big Data. So, I do not think it should be for everyone, but yes, the companies that want to use it and that need their own employees to use it or at least to make a step during the whole process of collecting data on analyzing data, they need to show their employees why it is important, and how their work is going to benefit the business as a whole. I think that is the main key point, which is not happening, as of today. And nowadays, I was saying you need to do this because I want and I do not need to give you any more explanation and if we make [all the employees] participants it will be a much easier process.

Ella Pommeranz: Like more involving everyone at all stages in the company.

Jorge Ferrándiz: Yes.

Ella Pommeranz: And then you also mentioned data privacy, like getting that trust when sharing your data. And how do you think we can ensure or built this trust, just like giving them statements to sign or I do not know. What are the steps to build trust between companies, in the case of sharing data?

Jorge Ferrándiz: Okay, at this stage there are like two kinds of trust: the trust with customers which you need to explain to them very clearly why are you asking them for the data and what are the purposes, […] which maybe is making personalized offers to them, or maybe it has nothing to do with commercial and it is only for own reporting in the business but you need to really clarify to the customers how the data is going to be used. And then the suppliers, your providers, need to know exactly why you want the data and sadly, you need to have strength with them. Maybe you are a small company asking your supplier for an integration with them to have the data collected automatically, not having to be writing all the data, you will face a problem which is that they will not listen to you. If you are a small company, most of the time they will not listen to you. And that is where HOSBEC gets in most of the time, because if we have 3, 4, 5 hotels that are interested in connecting with some CRM or reading system or whatever we go with the strength of 300 hotels. It is not the same as two individual hotels with 20 rooms each. So, we have a purpose in helping.

Ella Pommeranz: Okay, great I think this already like leads us nicely into the next question, which is about what data HOSBEC is using or like hotels in general. So, first you are collecting the data, then you are providing it to the 300 hotels that you are working with. So, what type of data does your organization have access to and use in its operations?

Jorge Ferrándiz: About the local data, we mainly work with occupancy rates as I showed you, but I want to show you like the whole process of the occupancy rates. Right before the pandemic, we were doing this automated but since COVID we wanted to have a weekly analysis of hotel occupancy so we changed our process to the following: We send them an online survey each week with the identification of the hotel, you can see the occupancy rate for the week, the nationalities of the occupancy rates of the week, the occupancy rate of the weekend only and on the books for the next week. We gather all this information, we then split it between markets, here we study Benidorm, Costa Blanca, Valencia and Castellón. We separate Costa Blanca and Castellón because we considered them strong enough to have their own analysis. But these are the four main markets we study. And in Valencia, not now because of the pandemic and we do not have enough samples. but before we were also studying Gandia [a city south of Valencia] as its own market. So, we gather all of the data for all the hotels, we get the means for each market and then with this information, we are publishing an interactive report, which is public for everyone. We upload this weekly […] with all the information for the actual year of 2022. For example, this last report was about the Easter holidays, so we have the week, the weekend and like the Easter period [=*Semana Santa*], which is from Thursday to Monday. And here we are studying the different nationalities, the evolution of occupancy rate for the whole week and we can understand [this] for each one of the markets. And, in addition to this interactive report, which is open to everyone, we also prepare a press note about the occupancy where we give our main conclusions.

Ella Pommeranz: And then you send that press report to all of the hotels, or just to the press?

Jorge Ferrándiz: We send this to all of our members. A hotel has like five or six contacts in our database: the director, the bosses of departments, so we send this to all of our contacts and to the interested press. Because every three months or so we have a communication sent to all of the press of the Valencian Community [telling] them if they want to receive our press notes, they can inscribe themselves in our mail chain or in our WhatsApp news channel that we have. Here you get the main conclusion of the report. In this case, it was a really, really high [occupancy]. But then we also do our analysis for each one of the markets, which is an analysis but it is basically these numbers turned into words. So, on the one side [we have] the occupancy rates that we study, our own data, because this data, as I showed you, is collected by us through online surveys. But we also do every three months, and every year an interactive report too. We also publish a big report called the "monitor" where we also introduce data from the national statistic institute. Here we make some of our data with some of the public data, the local data, as you mentioned here. For Spain, we do not study Spain as a whole because we do not [operate outside] the Valencian Community, but we study the overnights, the capacity, the open establishment, the employed people and the tourists from 2021. Although the main analysis is 2019 and 2021 because they are like the last normal years. The data from 2020 is not really useful. And for example here in Benidorm, because

we have a bigger sample done than the national [statistic] institute, we use our own occupancy rate for the year instead of the national one. But for the rest of the data, we use the national one. So here occupancy rate, both public and own data, and then we also have some reports, where we collaborate with them, but it was done by Visit Benidorm, which is the tourism office of Benidorm where they also study the capacity, online searches about the different markets. [He shows another page of the report] This is online flight searches and the evolution in-between years or by month.

Ella Pommeranz: And where are they getting the data from?

Jorge Ferrándiz: They are getting this from Mabrian. I was going to mention it on the part of the global data because we also work with some Mabrian data but yes, before we were using ForwardKeys. But Mabrian also expanded the products of information given and now with only Mabrian as a source we can do the same analysis. Okay, and I wanted to show you some of the data that we are collecting. This is one of the latest projects we are doing here at HOSBEC, which is Biontrend which is like an analysis tool for hotels. We are connected to the hotels' PMS, it is a property management system like the computer app that is at the reception that has all the bookings and the reservations, it has everything. We connect directly to the PMSs of the hotels and we get the information about occupancy rates, income, sold rooms, profitability, cancellations, everything. And doing it automatically, we allow the hotels to give us information without extra work which is, as I told you before one of the main problems. But here the problem is that we do not have all of our members inside the platform because it is like a really tough process to integrate. Some of the DNS providers do not help because they do not see any profit in the product being integrated with us. But right now we have almost 100 hotels in all of the Valencian Community, 60 or 70 from Benidorm so it is like a nice simple, but we want to make it bigger. And here, for example, I can see all the occupancy rates. I know this occupancy rate that I see is 100% real because it is not an online survey that they have to do manually. So I am 100% sure this data is real and I can compare as before by nationality, by booking site, if you only have breakfast or [full board], [or] by market, [which is] the same market as before: Benidorm, Costa Blanca, Valencia and Castellón. And here, for example, I can [compare] the data for this year April 2022 with 2019. So here I can see that in March we have been still 12% under what we used to be in 2019. But in April, as of now (as of April 20), we still have 10 days left, we are almost as we were in 2019, we are just 3% less. This is studying all of the markets, but if I only study Benidorm this may change. Here in Benidorm, for example, today April 20, 2022, I have a 32% occupancy on the books for September of this year, but on the same day of 2019 I had 31.9%, so it was less. I can say that today for September, we are better than in 2019 because we have more on the books. That can change with the months, but today we are better.

Ella Pommeranz: Amazing. And what do you say that the hotels are willingly sharing the data or are some a bit reluctant?

Jorge Ferrándiz: It was also some work at the start of the project because of the security issues. Many of them were not really comfortable giving all of this information. Because if it was just the occupancy rate okay, but we also have income, we have like millions of euros that the market is getting, the REVPAR [= Revenue per available room] which is used a lot in the hotel industry. And they were not really comfortable giving all of the information on the reservations. Because we also have a lot of information that is not so here, we have the nationality of the customers, if they come in family, if they come alone. Because all of this information will be used in the future for new analysis.

Ella Pommeranz: And would you say that they now start to see the benefit of sharing the data with others?

Jorge Ferrándiz: Yes. Especially now after the pandemic because during the pandemic it was a bit abandoned by the managers, because they did not have activity, so they did not have to use this. But now they are recovering, they are starting to use it much more and they are starting to see the real benefits of it. And this is making the other hotels, which were not inside the project want to get into this.

Ella Pommeranz: So, in the end, more people joined because of the pandemic?

Jorge Ferrándiz: Yes. This is also a program that is only open for hospitality associates, but we are planning to open this on a national level. It is just that HOSBEC cannot expand on the national level because they cannot get out of Valencian Community. But we are working with other organizations and associations that do work at the national level to expand this. Because we think it is a really powerful tool, and that it is getting really valuable data. Here, for example, I wanted to show you that we have all of the information about the bookings on the books of the hotels. Here, for example, I can see for each day, for example, yesterday the 19th of April, how many rooms were sold for each of the following months. So, I know that yesterday 12,000 rooms were sold inside of the Biontrend sample and I can see the level of the reservation written. It is really powerful and I wanted to show you that it is really new for us but it is gaining traction.

Ella Pommeranz: Yeah it is definitely amazing, and I hope that maybe in the future, more destinations can associate and produce this data as well. That would be amazing.

Jorge Ferrándiz: The problem with these kinds of tools which is the same that arose with the Valencian Intelligent Platform is that when an association or a local government develops a tool like this one, it is very limited to their own territory or to their own members. For example, here in HOSBEC, we have like public money to develop these tools but because it was done with public money, it can only be done to that public government which gave us the finances.

Ella Pommeranz: So, the Community of Valencia gave you funds for this project.

Jorge Ferrándiz: Yeah. HOSBEC works with a lot of public funds especially for the promotion of the destinations and for different projects that can profit the whole of the tourism industry of Valencia. Because it is the Valencian government giving the funds and not the national government, it is a bit difficult to expand internationally, but we are trying. And I was mentioning this because the same happens with the Valencia Intelligence Platform. Because it is a local government, the one who developed the platform. We do not have access, only local governments have access to that platform. So, for example, maybe the city council of Benidorm may have access, the city council of the city of Valencia or it is like destinations, who have access, but not companies or us as a hotel association. So, I cannot give you information about that because I have never seen it.

Ella Pommeranz: Well, the only thing I know that is publicly accessible is on their website, called the SIT, the tourism intelligence system. So, there is some kind of data that everyone can see. But the question is if everyone is using that or not.

Jorge Ferrándiz: Well, this is okay, but most of this data is extracted literally from the National Institute of Statistics. So, for example, we are already used to extracting this information from the National Institute of Statistics and we did not change our work process.

Ella Pommeranz: Ah okay, so you already had access to the same data before.

Jorge Ferrándiz: Yes. I think the only data that is not public already is this one about air traffic. But as we work with Visit Benidorm and Mabrian, we also have this information, the one that we consider useful from those sources. I did not know about the city of Valencia but we do not use it because we do not have reasons. Okay, and then about global data. The only global data that we use, I think it is the same as the city of Valencia, is the flight capacity. […] And we are planning to introduce different sources because we are trying to develop a predictive model and because of the predictive model we need to introduce as many sources as we can and then check which ones are useful for us or not, or which ones are related to the final occupancy rate or not. So here we are planning to introduce like flight bookings and online flight search. A lot of sources but I do not have all of them written down right now, but they are planning to introduce other sources here too. But on a daily basis, the only way we are using it is with the Big Data reports we do, which is a destination analysis. [This] is the Visit Benidorm reporting that we saw before. It is flight capacity, flight online search again, and here we have the evolution of bookings and cancellations by main countries. Here, we also have the social media sharing activity of the different markets, but this is mostly the job of Visit Benidorm, so I do not really know. I know they do a really nice job. But they are mainly focused on promotion, so that is social media, online search and different key indicators. Here, for example, the mean price of the different cities to see how it has been developing over the years. Here in the premises of Alicante, we used to have a really low hotel price, at least, but we have been also trying to increase it. At a slow pace, because if not the market will react negatively. But we are trying to increase the price. […]

Ella Pommeranz: And what type of data do you think will be necessary to have access to in the future, or is there something that is missing?

Jorge Ferrándiz: More than the type of data, what we are missing is mainly the part of the integrations with our different suppliers. Because we have a lot of information, but it is really difficult to make the business translate all that information that is presented in different papers or different online apps, it is very difficult to make the businesses gather all of the information and maybe give it to us. So, the fact of being integrated, for example in the case of Biontrend, with the proper PMS and getting all the information without any action from the hotel, it is what is giving us like real independency and real valuable information. So I think the trend in the future will be these automated processes and making the data collection tasks much easier and accessible for everyone.

Ella Pommeranz: Great. And then I think we can already move on to how you use all of the data that you collected. I think you already covered some of it. So, how would you say in general is Big Data used in HOSBEC, but also in hotels in general? Just a brief recap.

Jorge Ferrándiz: I can show you how the data is used in HOSBEC. We have occupancy rates which are public here in Costa Blanca since 2015. We have the public occupancy published periodically on our web if any person wants to see it. We also do press notes with a weekly analysis. […] We also have the Biontrend project, which is also a Big Data project, where we study not only hotel occupancy but overnights, sold rooms, income, profitability, etc, We also have the destination analysis, where we collaborate with Visit Benidorm. And I think that is like the only things that we do periodically, but we are also open to doing any analysis or any reporting that we consider good for the sector as long as we can, as long as we reach with the means we have.

Ella Pommeranz: Okay to maybe sum it up, in terms of basic analysis you are doing the reports, you have your dashboards and then you are also using the KPIs, right? Okay great. And then for advanced analysis, you already said that you are planning to introduce a predictive model, right?

Jorge Ferrándiz: Yeah. So, for HOSBEC as a whole, we are not planning on introducing advanced analysis. But inside the project Biontrend this is one of the developments planned for 2022 for this year. So we expect to have developed a predictive model, although we know a predictive model needs a lot of data in order to be useful. So this is the first year where we try to hit it off and give a lot of information and then next year we will try to just extract the sources that are really useful for us and that are really related to our final hotel occupancy which is the main KPI that we always use.

Ella Pommeranz: Great! Then let us move on to the potential barriers to Big Data analysis. And one of the ones that I found the most often when I did my research is actually the infrastructure. You need an internet connection, you need to have maybe sensors depending on what you are doing. So I was just wondering what type of infrastructure does your

organization currently have, but also, I want to talk about if you maybe know what kind of infrastructure do the hotels have?

Jorge Ferrándiz: Okay here in HOSBEC we mainly use cloud storage to store all of the data that we collect. For example, Biontrend is developed in [Microsoft] Azure, which is the Microsoft cloud. And then the normal ones are like Dropbox or Google Drive when it gets to storing information. But inside the hotels, I can tell you that there are a lot of apps or programs that are a must in the normal activity of a hotel. For example, I have shown you before the PMS, it is an informatics program that allows the hotel to be aware of the reservations that are done there, for the future. In addition to the PMS, a lot of hotels also have channel managers to connect with the different platforms that are selling the rooms, be it Booking, or their own website, they need a manager to funnel all of those sales. They also have most of the time a CRM, a customer relationship management system, which stores all the information of the customers, so they can identify repeated customers in the future, and they know that they like the room with two windows or they like the room with one additional bed. And they have prepared in advance, which gives a plus to the tourists to then come back to the hotel. So, it is mainly informatics programs or systems because internet connection I feel like it is something [being] ask[ed from] every organization, especially nowadays.

Ella Pommeranz: Okay, and is there anything that is needed for an optimal use and analysis of Big Data, now and in the future?

Jorge Ferrándiz: Something that I do not know if it is considered infrastructure, but it is something that is increasingly important in the sector, at least here in Benidorm and the Valencian community, is the knowledge of tools like Power BI. This was something that maybe five years ago no one knew what it was, or what it could be used for. And it is something that has been getting really important. […]

Ella Pommeranz: This is the perfect transition into the next section: data analysis skills. How would you describe the skills that employees in both HOSBEC and the hospitality industry currently have in regards to data analysis?

Jorge Ferrándiz: Here, I think it is shared between HOSBEC and the hotels because the main skills that are needed right now or are being used right now is on one side the analysis part, the reporting part, understanding the data and what data is giving us and showing us. And on the other hand, visualizing the data: making it neat and making it easy to analyze. So we have the one that reads the reports, for example, which may or may not have knowledge about analysis, which is the other hand. And in the future, I think the next step is having inside the organizations a more technical person, a person who can, for example, able to build up all these PowerBI schemes here with all of these analyses and all the integrations that we have. But maybe having in the organizations, both in HOSBEC and in hotels, a person capable of doing this, not only visualizing the data and analyzing it but building the

more technical part. And I think we will need to do that, I think it will be more normal to have a person with these skills in an organization, as the years go by.

Ella Pommeranz: Do you think every company is going to have at least one technical person? Also small- and medium-sized enterprises?

Jorge Ferrándiz: Yes. I think, in the case of SMEs possibly that person is not only technical maybe he is also part of the reporting and analyzing the data side, but something without at least minimal skills or knowledge about building and about tech, I think we will need that because right now always externalizing the technicians increases costs, reduces personalization, reduces if you want to change something from one day to another maybe you cannot.

Ella Pommeranz: Okay. What key factors do businesses like hotels or destinations need to optimize the implementation and utilization of Big Data?

Jorge Ferrándiz: Okay, I think I mentioned it before but until we have all the data collection processes under an automated system there is always a manual step that the human must do, and this is considered by many employees extra work and […] it is a tough work to make all the employees of a company understand the benefits of the Big Data or why we are collecting this data or why you have to do two steps more in your daily routine. I think what we need to optimize right now, or at least what we are doing here at HOSBEC with our hotel associates, is automating these processes, so they do not have to do anything. Because if we [give] extra work to a business that is, as you [saw] before, at an almost 100% occupancy rate, which has 30 people waiting at the reception, waiting for the receptionist to gather all the information of all of the tourists it is not working. So our main focus right now has to be on automating processes.

Ella Pommeranz: Okay, and again who do you think is going to initiate this change? Is HOSBEC helping hotels to do that, or should it be an internal change again?

Jorge Ferrándiz: HOSBEC is helping our hotels where HOSBEC can, which is, for example, in this project of Biontrend, or our weekly or even monthly occupancy rates collection. We are trying to make it as easy as possible. And the online survey that I showed you before, is something temporary. We started due to the pandemic because we wanted to have like a weekly analysis of hotel occupancy. But, once we are again [living in] normality we do with it all automatically we do not need anyone in the reception to do anything, to give us the data, because we are already connected to the PMS. HOSBEC can lead the change in HOSBEC-related tasks. […] There is a lot more to do inside hotels and it is the hotel direction, the hotel top level which, as we said before, proves the benefits of using Big Data in the organization and motivates all of the employees to work for it.

Ella Pommeranz: Great and then last but not least, how do you think will the adoption of Big Data in the tourism sector, and especially for hotels, evolve in the future?

Jorge Ferrándiz: I think related to this is the fact of integrating and automating all these processes will make it easier and more accessible for a lot more companies, and they will start to see the benefits of working with Big Data. For example, nowadays, if a small hotel wants to enter Biontrend or start giving us hotel occupancy in an automated way, maybe they need to wait one or two months until the supplier and us get everything around. But in the future, when this is much more settled down and much easier to integrate I think they will see the benefits like in one week or five days. And that will make other small companies become more interested too. Because big companies are already working with Big Data. We now need to focus on the SMEs.

Ella Pommeranz: I think there are a couple of steps to be taken and then, in the long run, I hope they are going to see the value of Big Data things. Okay, so that was already it, for now, thank you so much. Is there anything else that you would maybe like to add, but I have not asked about yet?

Jorge Ferrándiz: No, I just wanted you to have an idea of where our focus of Big Data here in HOSBEC is which are the public hotel occupancy rates that we publish on our website, the press notes that we release about different topics. The periodical one is about hotel occupancy but we may do it about any topic of interest. And the Biontrend project which is […] a really strong tool that I think something like that will be the norm in different sectors - having a collaborative tool where the different companies share the information that is then shown in an aggregated way, so they can take on decisions based on real information and not by their experience or by their impressions.

Transcript 7 – Joint interview with Carolina Navarro Correcher (Mobility, Cruises & City-Port Director at Fundación Valenciaport) and Joan Meseguer Llopis (Innovation & Port Cluster Development at Fundación Valenciaport)

Date: April 22nd, 2022

Duration: 1h 3min

Carolina Navarro: […] I am going to share the screen. […] Okay, just to start: HERIT-DATA is a project founded by the European Commission, through the Interreg Mediterranean Program. […] The aim of HERIT-DATA was to develop sustainable and responsible tourism management tools towards cultural heritage sites. We have been trying to implement some technology and innovation tools in order to better manage the flows of tourism. Pilots in HERIT-DATA, apart from Valencia, include some other pilots in France, Italy, Croatia, Bosnia and in Greece, as you can see here. All of them have applied the usage of Big Data for controlling their flows of tourism, using cameras or sensors with different

technology and the application of this technology is different from others, but the core is to use a cameras and sensors to monitor how people move around and in crowded cultural heritage sites. And, in particular, Valencia has been an example of using this technology, but for the cruise sector, that is a specific sector inside tourism as you know, and is quite different in some regards. All the results are available in our webpage that is [called] Snap4City that contains all the dashboards built by each of the pilot's site, including Valencia and including Valencia's dashboards. Thanks to the participation of the regional council of Valencia […] and the regional tourism government, we installed some sensors in the historical center at some key points, the most visited points, just to make some pilots. And in the entrance of the port. In this sense, we can detect with the sensors […] the mobile phones of the people that are around the sensor and detect when they move to another sensor, they can differentiate who has been entering by the port from the others. Thanks to this, we can differentiate the tourist flows from the general flow, who is a cruise passenger or not. Therefore, we are monitoring some potentially problematic sites in order to know what is the behavior of the cruise passengers. All this integration has been done by my colleague Joan.

Joan Meseguer: Yes, I just want to say that also apart from the sensors that we saw, in order to get a sense of the impact or environmental impact inside the buildings, because the environmental impact may damage the building inside and the heritage or protected objects. And then with a collaboration with Fundación Santa Maria La Real, which is another entity that was contributing to the project, they installed additional environmental sensors inside buildings also to gather these environmental impacts. Okay, so all these, together with the sensors that Carolina explained, it was integrated into the Snap4City platform. Because in the end, Snap4City has been the platform where you store the data and also where you represent the data, you visualize the data in a way that is appropriate for the objectives of the project, which were to know the situation in terms of occupancy and the crowdedness of places. So that people could choose whether to go to one site or the other, depending on how many people are in real-time at the current moment in those sites. This integration has been done partially here at the Port of Valencia foundation infrastructure, so we have some servers inside. And partially on the cloud, this is the Snap4City platform which allows you to manipulate data, so this is the one part that you can see here. […] So, if you want to we can go directly to the questions.

Ella Pommeranz: Yes, let us start with the more technical questions about the data availability. So, the first question would be what type of local or national data, so information collected in Valencia, does your organization have access to and use in its operations?

Joan Meseguer: So, about the national and local data we do use it because as mentioned by Carolina from one side, we use data collected by some sensors that have been installed by the regional council of Valencia. In the end that we had to kind of synchronize because they were going to do something similar, but not focusing on cruise passengers, just for general tourists or even people in general. So, then, we talked with them and said "Okay let us share our data. I want your data regarding the number of people that are in certain places. You

want my data regarding the number of people that are entering the city through the port". So what type of data do we get from them? Very raw data, so the number of people that are counted by the sensors in certain places. For instance, […] in the central market, […] and other places, like the City of Arts and Science etc. […] We also get data from the Port Community system, because the Port Community system is the data where the shipping agencies, like the ones for cruises, you know MSC and so on, so these are registering their future port calls, arriving to Valencia. This data is crucial for us because we do need to know when those cruises are going to come. Okay, so this is another platform [where] we are receiving data in real-time. We also started to check the National Institute of Statistics as well, there were some data there but this was like offline data. Just to get data on how many people were staying overnight in Valencia in the past year and so on, in the scope of the project, to get some KPIs. But this was the least fancy. Yes, but we also had to access this data. And if you want to ask about the Valencian platform, VLCi, yes, we do plan to have an integration with them. And we are in the process of integrating with them. Our intention is to push our data to them so that they can use these data to see how many people are in different places in real-time.

Carolina Navarro: We have collected extra information apart from the data that we already have as a port. In the case of Valencia, we have as Joan has explained, a Port Community System in which are the actors of the Port Community integrated, sharing data and information. That means we request the calls to send a list of passengers or all the data that has been communicated between the port authority and the other actors. That means shipping agents, the terminal. All these actors are integrated into our Port Community System that is called ValenciaportPCS. And it is quite important and well known in the whole world because we were pioneers in these contexts. This is our webpage. And also in ValenciaPort you can see the next calls.

Joan Meseguer: The thing here is that we connect through an API. This is like the front end, this is the user interface, but what we do for the project is connect our local platform, our infrastructure, and our software with the API that is provided by the Port Community System.

Carolina Navarro: As we are not users of the port, I do not have access. And we have a Port Community System in which shipping agencies, terminals, all the actors are involved, and they intend to interchange the communication electronically instead of on paper. And that was created 20 years ago, or more, and it has a quite long experience. But regarding the cruise sector, some information was missing on this platform, as it is not compulsory. And therefore within that data, we have also developed a management platform only for cruises. In this regard, we also have profiles for the terminal but also for the shipping agencies and the logistics company that provides the tour operators, for instance. Here we can also include information about the passes, transfers, excursions, all the information that is not compulsory on the Port Community Systems. Before any data, all this information was sent by email. And now we are working, it is still not put in production, but we are working on it, and all

this information will be available on a platform. You as an agent, as a tour operator will include your information about your tours for a specific cruise here, instead of sending emails. Therefore, we are collecting information, apart from the sensors, that is the project's focus on monitoring the tourist flows, where we are also improving the communication between the agents, and the stakeholders involved in this business. That is quite different from the business in other sectors. Therefore, these are the two main things regarding technology that we have developed within the project and with data. And the data that we already collected, we collect information that comes from the sensors and all the things that Joan has already explained, but we are also including all this logistic information regarding the cruise call.

Ella Pommeranz: Okay, and do you think there is any type of data that you are not using right now, but that might be necessary or useful in the future?

Carolina Navarro: Yes, we still need coordination with other actors and other actors need more information about us. For instance, we would like to know more information about some information that the cruise directly has. For instance, we do not know how many passengers are staying on board, how many passengers are hiring private excursions, how many are free. Therefore, there is still a lot to do in order to explore the behavior of the cruise passengers. Also, to measure the impact on the city. To monitor where they buy, how much they spend, what they like, what they propose for further trips.

Joan Meseguer: And also if you think about like the economical side of that. Just imagine that through the banks [like] Visa or MasterCard if they can share what is actually the economic impact of these cruise passengers. That would be great to see how we can optimize that. Or even improve the quality of the experience of the visit in Valencia.

Ella Pommeranz: And do you think these operators or actors are willing to share the data or do you think we are still reluctant?

Joan Meseguer: Yeah, as far as I know, you can get that data. But at the moment it is not free of charge, of course.

Ella Pommeranz: You got to pay for it.

Joan Meseguer: Yeah, so that is why in the end it is kind of a barrier. But the data is available because the banking sector is one of the most digitalized sectors in the industry but you know this is sensible data as well. This is another challenge. On the one side, some entities have data, but they are not digitalized, in the meaning that they are not open to the Internet where you can easily get data in real-time, so you need to digitalize. In order for us to get data for the project, we had to digitalize the port calling process of cruises or the management of cruises in the port. So this is the first challenge: digitalization. The second challenge is the exchange of data. You need to get into this GDPR law and all these barriers need to be solved as well. Sensible data and its exchange. And permissions, permissions to

install. If you want to install sensors you need to get permission from the council, from the owner of the building, if it is a heritage building: how are you going to deal with that. There are many things to take into account.

Ella Pommeranz: Do you think it makes sense to initiate the change from the national level so that the legislation has to be changed in order to enable the exchange of data or is that more like a local issue?

Joan Meseguer: That is a very good question because in some countries these legal things are managed locally and in other countries, they are in a more centralized way. So, well if it is local then it is easier because then you get involved, you have better contacts, better links to engage people and to get things done. But it is a major challenge.

Ella Pommeranz: In that sense, do you also think if the platform that Valencia has, if they would buy this credit card data from MasterCard would that be feasible? So they buy all the data and then provided it to tourism businesses and actors for free or maybe for like a small fee.

Joan Meseguer: Technically, it could be feasible. But as I said in terms of sensible data, I think this is a thing that needs to be checked. If you do it in bulk, let us say okay today in this museum it has been reported that 90% of all of the income was from abroad. You need to see the granularity of the data, because the more granularity you get, the bigger is the challenge to get that data. Because you dig into specific and sensible data. But it is something to be explored and actually the next step, we could take these economical side and see if it is possible. It is always possible but you need to solve these barriers step by step.

Ella Pommeranz: Next, I would move on to how Big Data is used in your organization. Just a short recap, because I think we already talked about it.

Joan Meseguer: We created a series of databases.

Carolina Navarro: Just a comment. We have been talking today about the use of Big Data in the tourism sector, but as a port, we are applying Big Data in a lot of use cases. And Joan is going to talk not only about a data platform collecting [data from] sensors, monitoring flows, but also, for instance, collecting emissions from the port. Because we are not from the tourism sector, we have a small lab that is dealing with that. But in our organization, we use Big Data for so many purposes in other projects.

Joan Meseguer: Good point. So, I would ask you, your question was targeting only tourists or more in general?

Ella Pommeranz: I think more about the cruise tourism part.

Joan Meseguer: Okay, so, in our case, we have, for instance, a NoSQL database [non-tabular database], which is getting data from the data that we were mentioning. We also have

PostgreSQL [= open-source database] to get data as well from the sensors. And this is combined and virtualized. And this is also scalable. Meaning that a city is virtualized, you can just scale it up and down according to your needs and the growth of the data. So basically we use virtualization and we use Docker [= software development tool for creating and running individual containers] and Kubernetes [= open-source system for operating containerized applications at scale] in order to easily scale the resources. So the more data you get, you need to scale up. We use integration mechanisms or tools such as Node-RED [= open-source flow-based tool and IoT platform and dashboard] to integrate data from other platforms and from sensors to store it in our infrastructure. We use APIs, we develop APIs in a scalable manner as well, and so that you can easily allow other platforms or other entities to get into this data in a fast way. We also re-use data from the regional council, which also uses platforms like Kafka, which is a data streaming database where you can store data in real-time and forward it to other platforms, users, etc. So we use this data, we use this platform […], but we do not have this platform on our premises. It is in the regional council. So in the end, […] the best here is the exchange of data, the integration with other entities that are doing something similar and we do not reinvent the wheel.

Ella Pommeranz: Okay, great and then once you collected all of this data, what type of basic analysis are you doing?

Joan Meseguer: We have Power BI tools as well, and we got Kibana [= proprietary data visualization dashboard software for Elasticsearch], Elastic Stack [= search engine] also to get data and visualize it through the Kibana dashboards. And the most important tool that we use to analyze to explore the data is Jupyter Notebook [= web-based interactive computing platform] and Pandas [= Python data analysis toolkit]. Jupyter Notebook is a tool, where you can get into the server that your data is in and you can develop or explore your data using Python, which is this programming language. So, you can use these Python libraries in order to manipulate and explore the data, present, plot data, and clean data, because cleaning the data is one of the most time-consuming tasks, so we are using these tools basically.

Ella Pommeranz: And then you have like the dashboards where you present your data, right?

Joan Meseguer: Exactly. Snap4City for the project. But for us, for mainly local purposes we use AngularJS [= JavaScript framework] to create dashboards or we use the ones provided by the Power BI tool or the Kibana.

Ella Pommeranz: And then are you doing some kind of advanced analysis in terms of predicting how many tourists are going to arrive in the future?

Joan Meseguer: This is the next step on our trip. Because we achieved the level where we get data, we store these data, we represent that data in real-time. But the next step, and we are going to do this in future projects, we need to look for funds also for this project, to

exactly use the data to predict how many people or cruisers are going to be in certain places in the future, in short- and medium-term. So that you can predict how overcrowded will this place be, which was the impact of the cruise, and in the future. But then you could link it to the economic impact as well. Which is going to be the economic impact? There are a lot of things that you can do from now on.

Ella Pommeranz: And what are some steps that have to be taken to get there?

Joan Meseguer: The next steps would be, I would say, to scale up. Well, I would install some more sensors. The more sensors you have, the more data you get where these passengers are moving in the city.

Carolina Navarro: Now it is a pilot and we have installed only a small quantity of sensors in some key points just to develop the algorithms in order to test and to differentiate [the flows of cruise passengers and locals]. We do not have a real picture of how they are moving. We are just testing the technology.

Joan Meseguer: And we tested that technology and we proved that it is possible. We developed this tourists differentiation mechanism to differentiate between cruise passengers and the rest. And now getting these data collected already, but enlarged with more data from other sensors, get it into these Jupyter Notebooks, using machine learning techniques, libraries to create machine learning pipeline models in order to predict the persons that will arrive at certain points in the city. And we check the impact, even the environmental impact inside the buildings to see if some of the buildings could be damaged potentially. In the long term, of course.

Ella Pommeranz: Great that sounds really interesting. Let us move on to the barriers of Big Data analysis. And one of the ones that I found the most often do my research was actually the infrastructure. And as we already talked about it, you need sensors, you need an Internet connection and I was wondering what type of infrastructure does your foundation currently have that enables the analysis of Big Data?

Joan Meseguer: First of all, you need to have servers. There are two trends here: you get the infrastructure locally. So, you enlarge your infrastructure with many more servers that you install on your premises and then you link all of them using, for instance, the tool Docker. And then you have a large pool of resources: CPU, RAM, etc. Because in the end when you do modeling, you do data exploration. This consumes a lot of resources. And the more you have the better. Right now, you have another trend, which is using cloud platforms. And we do use them. In this particular project HERIT-DATA, we just use Snap4City cloud but in general, we use other resources, other cloud platforms for other use cases. So the infrastructure we use is a combination of Bare-metal hardware, servers. And cloud services.

Ella Pommeranz: And what do you think about the trend that more and more businesses like destinations are outsourcing the infrastructure to external providers?

Joan Meseguer: So basically, I would say the trend is to move towards the cloud. So that you keep less Bare-metal in your premises because you need to maintain it […].

Ella Pommeranz: Okay, and what infrastructure would be needed for an optimal use of Big Data analysis, now and in the future? Is there something that is missing?

Joan Meseguer: I would say we are on a good way. I mean we use the proper tools at the moment. Here the most important thing is to stay updated on the technologies, and this is another challenge in IT.

Ella Pommeranz: Yeah, it always changes. So many new technologies.

Joan Meseguer: Exactly. In our team, we are updated on these things as far as I know. We are updated and we use the proper tools and the ones that I was mentioning to you to get Big Data. So, I would say that we are doing a combination of local infrastructure, meaning our hardware, and the cloud is good. Because sometimes you need to keep sensible data for the whole business data locally right, you cannot share it with others.

Ella Pommeranz: So, then, let us talk about the next challenge that I found which is data analysis skills. So how would you describe the skills that employees in the tourism industry or in your organization currently have?

Joan Meseguer: You have two things here. If you do it in your company, which I support that idea, the skills you need for these kinds are data scientists. [They need to be] familiar with these tools, the ones that I was mentioning like Python, Panda libraries, machine learning, and so on. And the other one is an expert in integration, developing integration mechanisms. Because in the end, your organization has data, your organization needs data. How are you going to exchange data? You need to develop some mechanisms to allow sharing this data. So, skills for developing these programs and data scientists.

Carolina Navarro: And I would like to add something that is quite important. In all the places, apart from ITC skills that I think are fundamental, I think it is also important to take into consideration that there are some experts in the operation that you are analyzing. Because someone has to define the KPIs, […] the input sources for this information, what outputs you need, what are the strategies that you are going to take into consideration, the results. We need good skills in IT and in data management, integration, all these computer skills that we have a very good team in our organization. But in order to go in the right direction, I think it is also important to have people with experience in the sector that you are trying to analyze.

Ella Pommeranz: And are there any skills that you would need that you do not have right now?

Carolina Navarro: Well, in my case I am not a data scientist and sometimes it is quite complex for me to understand the potentialities of Big Data for my sector. If I learn more about Big Data […] and if I have more experience of the potentialities and how I can use it for my sector, I think I will potentially know how to use it in the future. And I think also if the people that are good on Big Data and data know a little bit, not experts, but know a little bit about the sector, with both links I think it is better. Because we can create some innovation, innovative tools, with more experience than if you are on your own.

Joan Meseguer: Exactly and this is the thing. Here, the best is the combination of operational knowledge and technical knowledge. And that is why I support the idea of people that are working in the companies very focused on the maritime sector, to teach them, to participate in some workshops and to train them in order to get them better in these Big Data technologies. But it is not like necessary that you become an expert on machine learning.

Carolina Navarro: Exactly. Not about the details of technology, but just about the potentialities. How would you apply it and what can it provide to your sector.

Joan Meseguer: Where to look for it.

Carolina Navarro: To have a more strategic view, more than technical details.

Joan Meseguer: If I finally explore my data, which can be done by me in my company, great. I can provide the data and then all these data engineering. But at the end if I need to develop some models, perhaps I could do some simple models, but if there are complex models maybe it is better to subcontract that to some other company to do this.

Ella Pommeranz: Okay, good, and I would already move on to the last section. So what are the key factors that businesses and destinations need to optimize the implementation or the utilization of Big Data?

Carolina Navarro: I think that there is still a lot to do because we do not have so many Internet of Things tools in order to capture information. The problem is that there is a lot of information that is still not collected in a digitalized way.

Joan Meseguer: Digitalization, I would say. This is summarizing the key factor of the sector. Even museums, hotels, these touristic sites and so on, in order to get data from them, first you need to get digitalized.

Ella Pommeranz: I feel like a lot of them still do it manually like on paper.

Joan Meseguer: Sure.

Carolina Navarro: Exactly. With any data, we also try to define some KPIs. Also in other projects. Also when we did something related to the impact of the cruise sector on the city. We still do not have this information collected in that way. And some KPIs can be calculated

in [real] time because you collected information in a manual way but afterwards you cannot follow the progress of the KPIs because this information is not digitized.

Ella Pommeranz: I feel like once you have all of the information digitalized, then you can maybe also like automate it, and then we have this constant flow of information.

Joan Meseguer: Exactly.

Carolina Navarro: We are still [doing] the first steps. We are trying to go beyond that.

Joan Meseguer: As I said, other sectors are much faster than this one.

Carolina Navarro: Logistics and transportation are a very digitalized sector. But in the case of the tourism…

Ella Pommeranz: It takes its time.

Joan Meseguer: There is still a lot to do.

Ella Pommeranz: Okay, and then how do you think will the adoption of Big Data in the tourism sector evolve in the future?

Carolina Navarro: I think we will be able to collect a lot of information regarding how much money people spend and the preferences of the people, to evaluate this experience, to get feedback.

Joan Meseguer: Yeah, this combination of different angles: economical, societal, experience from sentiment analysis. In the end, economics is very important because you want to know which is the impact, you know how much employment tourism is generating. You want to know if these people [tourists] are providing money, I mean income to the local businesses. But also you need to know if the people that get into some restaurants think "Oh it is too expensive in this place", and this has a negative impact on the experiences. All these things need to be connected somehow and to have an overall view of everything. This situational awareness in this area needs to be taken in a more global way. Getting different factors and different links.

Ella Pommeranz: Great and then now linking it back, because I think what you mentioned are actually some of the opportunities of Big Data right, so I would jump back. So could you please quickly sum up what are some of the main opportunities of Big Data for the tourism sector?

Joan Meseguer: Basically, I would say to get a clear view or improve this situational awareness regarding the impact of tourism in a locally and globally, but mainly locally, way and for the different sectors. So, for instance, in our case, cruise tourism. So, to see the impact, you need to have to be aware of the situation in real-time. And in the future, of

course, if you could advance with Big Data it allows you to create the tools, to make or apply preventive actions and counteractions to some problems that may occur. For instance, a place could potentially be overcrowded so thanks to the Big Data you can advance to it and apply some rules that overcome the situations. And Big Data allows us on one side to protect the local heritage and also from the other side for tourists to improve their experience [during] their visit.

Carolina Navarro: I think there are these two directions. For tourism destinations managers, they can improve their knowledge about the tourism sector. How the behavior, the experience, everything, the impact. There are a lot of fields that can be explored by Big Data and they can use all this information for designing new strategies and strategic tools. From the tourist's experience, you can also offer a better experience in your destination with the application of Big Data. To sell them new ways of discovering their destination, new alternatives, how to manage all the things that they need for the trip, booking different types of experiences, tours, transportation, hotels. Everything can be integrated in another way. And I think it is a new way for tourists to discover the city. With some virtual experiences, or with apps for discovering the city, augmented reality. There are so many IT tools that can be implemented towards improving the experience of the tourist.

Ella Pommeranz: Very good. And then I would move on to more like the challenges. So I think as we all know by now, a crisis can hit us anytime like the pandemic or the war. So how do you think can Big Data help us to cope with these challenges, coming from the broader environment?

Carolina Navarro: In the case of the pandemic in the cruise sector, for instance, you can use Big Data to monitor the people. We do not have this information, but Big Data can support, for instance, the collection of information about which places have been visited, if you have been in contact with other people, if you are in contact with someone that has been having COVID.

Joan Meseguer: For some restrictions like the number of people allowed in a museum or any closed space.

Carolina Navarro: Counting people inside in order to prevent that there are more people than planned. There are a lot of things, we have colleagues that are working on pandemics. And some projects related to helping you manage pandemics.

Joan Meseguer: The Toscana region, I think, was analyzing this in the project actually. Thanks to these Big Data sentiment analyses and so on, and the sensors to see the things that were mentioned by Carolina.

Carolina Navarro: Joan and I are not experts on this. We are doing things for ferry passengers, yes, for instance, or also for cruise passengers, but we are still working on it.

Joan Meseguer: And how did the pandemic impact our projects? Well, during the pandemic cruise ships were not approaching or arriving at the port, so for some time, we could not collect valuable data. The pandemic is an atypical situation which, even if you have sensors installed and collecting data, should not be treated as representative data. So, these kinds of pandemics and even war, people stop traveling and so on, this might also introduce some noise to your data compared to the past, in normal conditions.

Ella Pommeranz: And then the very last question. You already talked about two challenges which were digitalization and the exchange of data, so like the sensitivity of data. Can you think of any more main challenges regarding Big Data?

Joan Meseguer: Related to Big Data, well from the IoT side of the Big Data I would say the permissions to install new sensors in some places. For instance, sometimes you need to contact the council and sometimes you need to contact a private entity managing the touristic side. The port authority because you want to install a sensor in some of the terminals. So you need to be in contact with many organizations and to get approval to do this. And also a very challenging thing, which looks simple but it was a challenge for us, is if you want to install sensors in some places you rely on a power source. And, and if some people are managing this power, sometimes the sensor stops working and you do not know why, and all these things need to be taken into account as well to ensure that you have continuous power on your sensor. And the environmental conditions. It is not the same installing a sensor nearby the sea or in the city center. You know, the winds, rain, and all this might also damage your sensor.

Ella Pommeranz: Anything else you want to add Carolina?

Carolina Navarro: No, I think is good okay.

Ella Pommeranz: Good because I think from my side that would be everything for now. Thank you so much for your questions and your insights I really appreciate it and have a nice day.