Information Technology & Tourism (2026) 28:26
https://doi.org/10.1007/s40558-026-00368-0

ORIGINAL RESEARCH

# Innovating data-driven tourism reputation management: methodological foundations of the tourism online reputation index (TORI)

**Julio César Puche-Regaliza[1] · Isabel Marcilla-Lombraña[1] · Paula Antón-Maraña[1] · Pablo Arranz-Val[1]**

Received: 9 June 2025 / Revised: 24 February 2026 / Accepted: 2 March 2026
© The Author(s) 2026

## Abstract

The widespread use of social media has significantly amplified the role of online reputations in shaping the image and competitiveness of tourism destinations. This study proposes an innovative methodology that combines big data techniques with geolocated user-generated content to develop a comprehensive tourism online reputation index (TORI). The TORI aims to quantify and monitor tourists' perceptions of destinations in a structured and scalable way. The methodology integrates the cross-industry standard process for data mining (CRISP-DM) and knowledge discovery in databases (KDD) frameworks to ensure a rigorous, systematic approach to data collection, processing, and analysis. An ontology is developed to categorize and structure the diverse attraction points within destinations, and natural language processing (NLP) techniques are employed to perform sentiment analysis and generate tourist profiles on the basis of online reviews. The proposed methodology is validated through a case study in the province of Burgos, Spain, illustrating its practical relevance for enhancing data-driven decision-making in the context of smart tourism destinations (STDs). The results are presented through an interactive scorecard that facilitates intuitive interpretation by tourism stakeholders and supports strategic planning. From a theoretical perspective, this study contributes to the literature by offering a quantitative and standardized approach to measuring online reputation, addressing the lack of integrated tools and human-centered vision in current tourism research. In practice, it provides a replicable and adaptable solution for destination managers, particularly in rural and sparsely populated areas, to improve reputation management, support sustainable development, and strengthen destination competitiveness in the digital era.

**Keywords** Smart tourism destinations · Tourism intelligence system · Inland rural tourism · Synthetic tourism index · Tourism decision making · Sustainable tourism development

Extended author information available on the last page of the article

## 1 Introduction

The growing influence of user-generated content on digital platforms has reshaped how tourism destinations are perceived, promoted, and managed. Online reputation, reflected through tourist reviews, ratings, and comments, has become a critical element of competitiveness, influencing tourist choices and public perceptions. Xiang et al. (2017) and Buhalis and Sinarta (2019) highlighted how digital interactions shape destination image, whereas Marine-Roig and Huertas (2020), Martín-Galán (2020), and Wu et al. (2024) emphasized their role in managing competitiveness and reputation. In addition, research on digital word-of-mouth plays a central role in tourists' decision-making and in shaping the perceived image of destinations (Zeng and Gerritsen 2014; Filieri et al. 2015). Collectively, these studies underscore the growing relevance of online reputation as both a driver of competitiveness and a tool for strategic management in tourism.

This relevance is particularly evident in the framework of smart tourism destinations (STDs), where digital interaction and feedback are core components of governance and competitiveness. The integration of information and communication technologies (ICTs) opens new possibilities for data-informed management, service personalization, and long-term sustainability in tourism. Previous studies have shown how ICTs facilitate new service models and management strategies (Lee et al. 2013; Li et al. 2017; Mandić and Garbin-Praničević 2019). Recent research has emphasized their transformative role in enhancing tourist experiences and destination competitiveness (Grundner and Neuhofer 2021; Samala et al. 2022; Lasisi et al. 2023). Moreover, ICTs are increasingly recognized as enablers of sustainability and inclusive development within STDs (Sustacha et al. 2023; Alsharif et al. 2024; Wei et al. 2024; Aransyah et al. 2025).

The STD concept, promoted by the European Commission and operationalized in Spain by the State Society for the Management of Innovation and Technologies (SEGITTUR), provides a reference framework for innovation, accessibility, sustainability, and technology-led governance (SEGITTUR 2020; DTI 2021; STD 2024). A recent application of the SEGITTUR model can be found in Aïdi and Fabry (2024).

Within this framework, particular attention is given to the technological axis, which includes indicators such as T15 (*Social Media Analytics*), T6 (*Indicator Scorecards*), and T7 (*Tourist Characterization Tools*) (UNE 2022). These indicators measure the use of digital tools to monitor online presence, evaluate destination performance, and understand tourist behavior. The methodology proposed in this paper contributes directly to the improvement of these indicators by offering mechanisms to (a) measure online sentiment, (b) monitor supply and demand through data scorecards, and (c) extract insights into tourist preferences and experiences.

Despite the strategic importance of online reputation, the tourism sector still lacks standardized, data-driven tools capable of monitoring and quantifying this dimension across all attraction points within a destination (Moro et al. 2016; Vianello et al. 2023). The existing approaches are often fragmented, platform specific, or limited to individual assets, making it difficult for managers and policymakers to obtain a comprehensive view of visitor experiences. This gap limits the ability of destinations to translate the wealth of digital feedback into actionable strategies for competitiveness, sustainability, and governance.

To address this gap, the main objective of this study is to design a systematic and innovative methodology for constructing a synthetic tourism online reputation index (TORI). The index integrates big data techniques and geolocated information to capture the perceptions of tourists regarding destination experiences on the basis of online ratings, reviews, and sentiment expressed across digital channels.

The methodology also generates several outputs: (1) the development of an ontology to categorize and analyze the tourism supply across attraction points; (2) the application of natural language processing (NLP) to perform sentiment analysis and tourist profiling; and (3) the design of an interactive scorecard to visualize key indicators (specifically, the index generated upon achieving the main objective of the study), thereby facilitating strategic decision-making by tourism stakeholders. By combining these elements, the study provides a scalable and replicable approach for assessing online reputation at the destination level on the basis of all its attraction points.

Spain plays a leading role in global STD research and implementation, with numerous initiatives and academic studies contributing to the development of conceptual models, diagnostic tools, and strategic frameworks (Shafiee et al. 2021, 2022; Ercan 2023; Hiererra et al. 2023; Liu et al. 2023; Absari et al. 2024; Herrera-Prado et al. 2024). The methodology proposed in this study builds on and contributes to this body of work by offering a data-driven approach specifically focused on online reputation as a key dimension of destination competitiveness and sustainability.

The structure of the paper is as follows. Section 2 presents the related literature and situates the study within current academic debates. Section 3 details the proposed methodology, including the development of the TORI index. Section 4 discusses the application of the methodology to a case study in Burgos, Spain, illustrating its practical relevance. Finally, Sect. 5 concludes with key insights and recommendations for future research and destination management.

## 2 Background and contribution

In this section, the problem-specific literature related to this study is presented. The theoretical and conceptual frameworks are analyzed. From the literature review, key gaps emerge, highlighting the main contributions of this work.

### 2.1 Theoretical framework

The theoretical foundations of this study draw on three interconnected strands of literature: (i) STDs, (ii) technological tools and data analytics in tourism, and (iii) the use of online reputation indices. Together, these elements support the conceptual and methodological development of the tourism online reputation index (TORI), which integrates quantitative and qualitative data, applies a human-centered perspective, and aligns with the technological axis of the STD model.

### 2.1.1 STDs and cocreation

The evolution of smart destinations provides a conceptual basis for understanding the central role of ICTs and cocreation in modern tourism. During much of the 20th century, the value of a destination was defined mainly by its geographical features, entertainment and accommodations, and transport infrastructure. Visitors were regarded as passive consumers, with little influence on the structure of the destination (Howie 2003; Jovicic 2019). By the end of the century, however, a new perspective emerged: destinations were increasingly understood as networks of interconnected public and private organizations, where multiple stakeholders shape the tourism supply (Baggio and Cooper 2010; Baggio and Sainaghi 2011; Ivars-Baidal et al. 2024).

The rapid development of ICTs accelerated this transformation, and smartness began to play a role in tourism destinations, leading to the concept of STDs (Bastidas-Manzano et al. 2021). ICTs enable the immediate exchange of information and knowledge, supporting knowledge-based destinations in which cooperation and interaction become central. In this approach, cooperation plays a key role, as does the tourist, whose actions can influence the structure of the destination. Therefore, through two-way communication processes—especially via social media—visitors can influence destination structures and contribute to competitiveness. This role of cocreation has been highlighted in foundational studies on smart tourism (Buhalis and Amaranggana 2015; Hunter et al. 2015; Xiang et al. 2015), as well as in more recent analyses that emphasize digital interaction and user participation (Mohammadi et al. 2021; Carvalho and Alves 2023; Garanti 2023; John and Supramaniam 2024).

This cocreation process reinforces the technological axis of the STD framework, highlighting the importance of digital interaction and visitor feedback. Following these ideas, smart tourism thus moves beyond simply "visiting" places to generating experiences mediated by ICTs. As Gelter et al. (2021) stated, ICTs represent a competitive advantage for the tourism sector as a whole and, particularly, for STDs. The development of cutting-edge technologies, combined with the previously mentioned dynamic codesign processes of destinations, contributes significantly to enhancing the competitiveness of STDs.

Beyond competitiveness, sustainability has become a central requirement for contemporary tourism development. ICTs contribute to this objective by supporting cocreation within smart ecosystems, thereby promoting the sustainability of STDs. Early contributions emphasized the potential of ICTs to integrate sustainability into smart ecosystems (Ivars-Baidal et al. 2017; Ribes and Baidal 2018), while more recent studies have reinforced this link by exploring how digital tools foster inclusive and sustainable development (Díaz et al. 2023; Borges-Tiago and Avelar 2025). Tourist destinations have undertaken smart initiatives to improve the quality of life of residents and promote their long-term viability. This has led to the recognition of the concept of sustainable smart tourism development as a significant area of research interest (Cavalheiro et al. 2020, 2021). Gretzel et al. (2015) defined it as tourism supported by the endeavors to obtain data with the purpose of later transforming it, using developed ICTs, into online experiences and business value propositions, with the ultimate objective of creating an efficient and sustainable destination that enriches the experiences of every stakeholder.

Recent studies have begun to explore the interdependence between sustainable development and smart tourism destinations (Shafiee et al. 2019), although this research is still in its early stages. The suggested directions include expanding the concept of STDs to encompass smart tourism cities within broader geographical areas and adapting the smart approach to local and sparsely populated territories to ensure social inclusion (Buhalis et al. 2023). The lack of inclusion is particularly relevant in rural areas, where digital literacy tends to be low and traditional small businesses often remain outside the smart ecosystem (Ballina 2022; O'Connor 2023). This raises concerns about uneven access to the benefits of smart tourism and the risk of widening digital divides.

Complementing these discussions, Gomes et al. (2024) identified five interconnected dimensions that shape tourism and hospitality research: knowledge transfer in tourism, networking and innovation, sources of innovation, smart tourism ecosystems, and innovation research in tourism. They concluded that tourism development should be approached as a regional competence, grounded in strategic networking and the externalization of regional knowledge flows.

Taken together, these studies show that the smart destination paradigm depends not only on infrastructure and innovation but also on the strategic use of real-time data—particularly tourists' opinions and digital interactions—to guide inclusive and sustainable development. This provides a direct link to the analysis of online reputation and to the need for robust metrics that synthesize tourists' perceptions at the destination level, which is where this work makes its primary contribution.

#### 2.1.2 Technological tools and data analytics

Building on the previous conceptual evolution, a second strand of literature emphasizes the technological transformation of the tourism sector. The relationship between technology and tourism has been widely explored, especially in terms of its impact on destination management and the tourist experience. For example, Berné-Manero et al. (2011) highlighted how technologies enhance management efficiency and visitor experiences, while more recent studies have described the shift toward Tourism 4.0, where digital and physical platforms are increasingly interconnected (Peceny et al. 2019; Kalandarovna-Abdurakhmanova et al. 2022).

Tourism 4.0 represents a new value ecosystem based on high-technology services. It is characterized by interoperability, virtualization, decentralization, real-time data gathering and analysis capability, service orientation, and modularity (Pencarelli 2020). According to Herrera-Prado et al. (2024), technology has therefore become a key facilitator of trip planning and execution, enabling destinations to rely on intelligent techniques for efficient management.

Consequently, smart technological tools such as smart sensor networks, open data, cloud computing, geopositioning systems, blockchain technology, big data, data mining, end-user devices, services and applications, monitoring and information systems, Internet of Things (IoT), radio frequency identification (RFID), near field communication (NFC), 5th generation mobile network (5G), artificial intelligence

(AI), augmented reality (AR), or virtual reality (VR), with different purposes have emerged (Sigala 2018; Buhalis et al. 2019; Aliya et al. 2023; Cheng et al. 2023; Fauzel et al. 2024).

Within this context, a wide array of smart technological tools has emerged to support tourism development, ranging from infrastructure-oriented systems (e.g., smart sensor networks, monitoring and information systems, cloud computing, geopositioning) to advanced digital solutions (e.g., blockchain, big data, data mining, Internet of Things (IoT), radio frequency identification (RFID), near-field communication (NFC) and user-focused applications (e.g., artificial intelligence (AI), augmented reality (AR), virtual reality (VR), end-user devices, services and apps). These tools, which have been discussed in studies such as Sigala (2018), Buhalis et al. (2019), Aliya et al. (2023), Cheng et al. (2023), and Fauzel et al. (2024), illustrate how technology-driven innovation is reshaping tourism ecosystems.

Among these smart technological tools, four combinations have recently gained particular prominence: (1) the use of online platforms for collecting spatial data; (2) statistical and AI tools to analyze the collected data (e.g., identifying tourist behavior and satisfaction, patterns and trends and developing personalization and recommender systems, prediction and forecasting systems, or natural language processing (NLP) systems); (3) geolocated information and geographic information systems (GIS) for visualizing spatial data; and (4) big data to achieve a better understanding of tourism demand, tourist supply, and other tourism issues. This trend has been supported by a wide range of studies, from early research on ICT applications in tourism (Januszewska et al. 2015; Li et al. 2018) to later analyses of big data and AI for tourism analytics (Bulchand-Gidumal 2020; Sunagar et al. 2020; Chen et al. 2021; Li et al. 2021) and more recent work using advanced techniques such as deep learning and geoanalytics (Hu et al. 2022; Lyu et al. 2022; Mariani and Baggio 2022; Rong et al. 2024; Wu et al. 2025). These developments offer new challenges and opportunities for the sector, in addition to the adaptability previously mentioned.

A growing body of empirical research has illustrated the potential of these tools. For instance, Vecchio et al. (2018) demonstrated the use of big data, social media platforms (Facebook, Twitter, Instagram), and AI tools (Keyhole, Buzztrack) to support decision-making among STDs. Studies based on geotagged data and machine learning models have identified destination dynamics and the factors influencing tourists' choices, such as that of Giglio et al. (2019), who used Flickr photos. Forecasting frameworks have also been developed, as in Sun et al. (2019), who combined machine learning with internet search indices from Google and Baidu to predict tourist arrivals. Other studies have explored deep learning and social geodata to map spatial, temporal, and demographic tourist flows (Paolanti et al. 2021) or to reveal areas of interest and spatial imbalances through location-based networks such as Twitter, Foursquare, and TripAdvisor (Nolasco-Cirugeda et al. 2022). Finally, Leiras and Eusébio (2023) analyzed accessibility and visitor satisfaction by applying text-mining techniques to Google Maps reviews.

However, the expansion of smart technological tools has also increased the complexity of destination image formation (Ghazali and Cai 2013; Wang et al. 2021; Díaz-Pacheco et al. 2024). This evolving landscape not only reshapes how destinations are perceived but also directly influences tourist behavior. Foundational studies have stressed the role of ICTs in shaping tourist choices (Buhalis and Law 2008; Hung et al. 2011), whereas other studies have confirmed that technology-driven interactions significantly impact competitiveness and visitor decisions (Tavitiyaman et al. 2021; Chia et al. 2021; Hamdy et al. 2024).

### 2.1.3 Online reputation and reputation indices

Tourist ratings and reviews shared through social media and digital platforms also play a decisive role in shaping destination image and influencing the decision-making process of prospective visitors. Early work has reported that reviews help reduce uncertainty in destination choice (Saranow 2004), and subsequent studies have consistently confirmed this effect (Lin and Huang 2006; Wenger 2008; Bosangit and Mena 2009; Tham et al. 2019). In this context, Pan et al. (2007) demonstrated their usefulness for identifying the strengths and weaknesses of destinations, whereas Pan et al. (2021) analyzed how the images perceived by members of a tourist's social network affect destination choice. Beyond image formation, reviews can also serve as proxies for real tourist satisfaction, supporting continuous improvement processes and the design of loyalty strategies (López de Ávila 2015).

Against this backdrop, the concept of online reputation management has gained increasing relevance, understood as the perception and evaluation that tourists form about a destination and reflecting its prestige in the digital environment (Meijomil 2022). This relevance requires destinations to monitor and analyze online presence and visitor feedback across social media, travel blogs, and digital platforms. While online reputation is not an ICT in itself, its management depends heavily on ICT-based tools. One of the most widely used techniques is sentiment analysis, which is often conducted by applying NLP algorithms to tourist-generated content at different stages of the travel cycle. It has been used to link ratings with sentiment in reviews (Geetha et al. 2017), to map the methodological landscape and operational pipelines (Álvarez-Carmona et al. 2022; Xu and Lv 2022), and to advance practical applications such as improving tourism communication and domain-specific insights (Alhajri 2024; George and Ramos 2024).

Given this situation and despite the clear technology-driven evolution of tourism, many approaches still lack a human-centered design. Scholars have stressed the need for smart tools that acknowledge the central role of people and prioritize their well-being in tourism (Stankov and Gretzel 2020; Gani et al. 2024).

Building on the discussion of online reviews and sentiment analysis, the third strand of literature concerns the development of online reputation indices in the tourism sector. This section examines the existing indices, their data sources, methodological approaches, and current limitations.

In this context, the connection between online reputation management and social media platforms (such as blogs, networking sites, photo and video stories, chats, and forums) has become essential. It is increasingly understood as a process of monitoring and interpreting user-generated content, particularly ratings and reviews, to influence and shape destination image (Kushcheva and Eilola 2023).

Going one step further, research into online reputation indices within the tourism sector has evolved significantly, focusing on how digital tools and data analysis shape the perceptions and management of destinations. For example, Platov et al. (2020) outlined a framework for analyzing destination reputation in online environments, identifying key components, objects of analysis as indicators of reputation, and technologies involved. However, they did not provide a methodology to quantitatively define an index, noting the need for further empirical research to determine the role, significance, and relative weight of each individual component.

Building on earlier work such as the tourism online reputation (TOR) score (Pollák et al. 2016), Cillo et al. (2021) examined the role of big data analytics in strengthening the online reputation of niche destinations. Their findings underscored the strategic value of digital reputation management, but the study itself remained largely qualitative, leaving quantitative validation an open research challenge.

To address the limitations of earlier qualitative approaches, several studies have advanced quantitative methodologies for defining online reputation indices. For instance, Carrasco-Santos et al. (2021) created a generic index using TripAdvisor ratings (Likert scale, 1–5) to reflect visitor experiences at individual attractions in Marbella. Their work stands out for moving beyond destination-level assessments and focusing on the reputation of specific points of interest. Similarly, Pártlová et al. (2022) assessed the reputation of 13 European tourist cities by analyzing content across Instagram, YouTube, and Twitter. They developed an index based on posts/media, comments, likes/views, and followers/subscribers, concluding that a greater volume of content does not necessarily translate into stronger user engagement.

Other contributions have explored more sophisticated analytical tools. Qin et al. (2022) proposed a recommender system that ranks attractions using aspect-level sentiment analysis combined with multicriteria decision-making under fuzzy conditions. Bui et al. (2021) presented a structural model integrating constructs of a destination image with computational methods (NLP and multimedia analysis), while recognizing the challenge of accurately capturing complex human emotions through photos and reviews. Finally, Antón-Maraña et al. (2023) designed a tourist reputation index based on big data and NLP techniques and calculated a weighted average of positive, negative, and neutral comments. A key limitation of their approach is the reliance on textual comments exclusively, without integrating numerical ratings, which could provide a more standardized measure of reputation.

To summarize, the current online reputation indices in tourism rely on diverse data sources—ratings, number of reviews, comments for sentiment analysis, and increasingly multimedia content such as videos and images extracted from social media platforms. Despite these advances, three main shortcomings persist. First, many indices still privilege a qualitative approach rather than a quantitative approach. Second, there is little homogenization or standardization in the evaluation of tourist destinations, which hinders comparability. Third, most indices focus on individual attractions, neglecting the integrated perspective of the destination as a whole and the collective experience generated by its different components.

To better contextualize these limitations and reinforce the fragmented nature of the existing research, Table 1 provides a comparative overview of the previous studies. It contrasts with their data sources, analytical techniques, type of index, and methodological limitations.

**Table 1** Previous studies on online reputation indices

| Study | Data sources | Methodology | Type of index | Limitations |
|---|---|---|---|---|
| Carrasco-Santos et al. (2021) | TripAdvisor ratings (1–5 Likert scale) | Index based on average ratings of individual attractions | Simple arithmetic average | Focused only on individual attractions; lacks destination-level aggregation |
| Pártlová et al. (2022) | Instagram, YouTube, Twitter | Interaction metrics (likes, comments, followers, views) | Activity-based indicator | Quantity ≠ quality of interaction; no sentiment or content-based analysis |
| Qin et al. (2022) | Online reviews | Aspect-based sentiment analysis + MCDM (fuzzy logic) | Recommender system | No synthetic or destination-wide metric; limited generalization; complex methodology |
| Bui et al. (2021) | Photos, comments, reviews | Structural model + NLP + multimedia content analysis | Construct-based multidimensional index | Difficulty capturing nuances of human emotions; lacks a unified composite index |
| Antón-Maraña et al. (2023) | Tourist comments in Twitter | NLP + weighted average of sentiment (positive/neutral/negative) | Weighted sentiment index | Ignores numerical ratings; lacks standardization/homogenization across destinations |
| This study (TORI) | Google Maps (ratings + reviews) | NLP + polarity + thematic ontology | Composite, geolocated, weighted index | Initial application to a single case study; scalability to be tested; potential imbalance in weights |

This synthesis underscores the persistent gap in the literature and justifies the development of a standardized, scalable, and human-centered index—such as the one proposed in this study—that integrates quantitative ratings and sentiment analysis at the destination level.

## 2.2 Conceptual framework

As discussed in the previous subsection, the development of the TORI is grounded in an integrated conceptual framework that connects three key strands of the literature—STDs, the use of technological tools and data analytics, and the measurement of online reputation. These components form a coherent structure that not only justifies the design of the index but also positions it within the broader discourse on smart and sustainable tourism.

First, the STD paradigm serves as the foundational model, highlighting the need for innovative and participatory governance in tourism destinations, supported by real-time data and digital interaction. Second, the integration of smart technologies, particularly big data techniques, NLP techniques, and geolocated information, provides the methodological basis to capture, analyze and visualize large volumes of user-generated content in a way that is both scalable and replicable. Finally, the third strand involves the analysis and quantification of online reputation through tourist-generated ratings, reviews, and sentiments, which act as proxies for tourist satisfaction and cocreation of the tourist experience.

Figure 1 presents a visual synthesis of this framework, illustrating how these components converge in the construction of TORI. By combining structured data (e.g., ratings and review counts) with unstructured data (e.g., textual reviews analyzed through NLP) and linking them to geospatial information, the TORI offers a comprehensive, standardized, and human-centered tool to evaluate the online reputation of an entire tourist destination, not only attraction points.

This framework thus aligns with the technological axis of the STD model, promotes strategic decision-making for inclusive and sustainable tourism development, and serves as the foundation for the methodological development presented in Sect. 3, particularly in the construction and validation of the TORI.

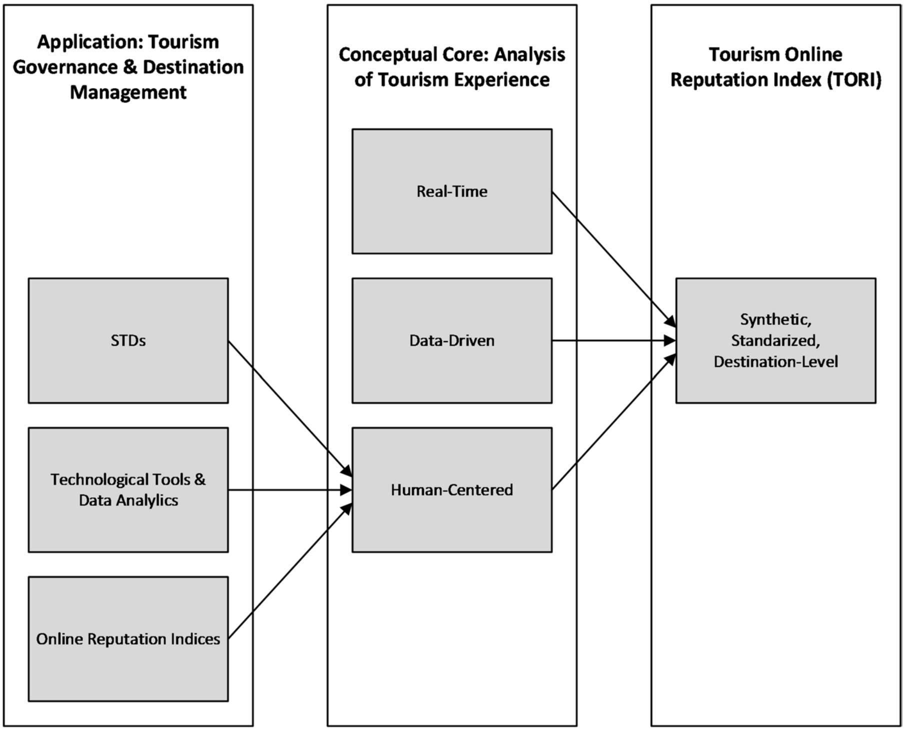


**Fig. 1** Conceptual framework. Prepared by the authors

### 2.3 Contributions

The gaps identified in the literature review are addressed through several contributions that extend the study's value beyond the immediate case.

First, the main contribution is the design of a systematic and innovative methodology that integrates big data techniques and geolocated information for the development of a synthetic index of online reputation (TORI). Unlike existing approaches, which often focus on isolated assets, the TORI provides a quantitative and standardized metric covering the entire set of attraction points within a destination. This territorial approach enables the evaluation of inland, rural, local, and sparsely populated areas, thus supporting regional competitiveness, promoting strategic networking, and advancing the social inclusion of disadvantaged areas. Methodologically, this represents a scalable framework that can be replicated in other territorial contexts.

Second, the study contributes to conceptual and theoretical advances through the development of an ontology that categorizes attraction points. This ontology offers a structured way to analyze destinations, enabling the construction of both global and partial indices by category. By doing so, the research provides a transferable framework that supports comparative studies, benchmarking, and theory building on how online reputation relates to different components of the tourism system.

Third, the integration of tourist-generated ratings and reviews, the review volume, and sentiment analysis via NLP introduces a human-centered perspective. This contribution emphasizes cocreation, reflecting tourists' roles as active agents whose digital interactions inform destination strategies in real time. Theoretically, this approach contributes to debates on participatory governance in tourism and the growing importance of visitor feedback as a resource for decision-making.

Finally, the introduction of an interactive scorecard to monitor supply- and demand-related information provides a concrete managerial tool. This practical contribution strengthens the technological axis of the STD model proposed by SEGITTUR, facilitating more transparent, data-driven, and sustainable governance. By aligning with the STD framework, this research offers both policymakers and practitioners a model that can be applied to achieve higher levels of innovation, competitiveness, and sustainability in tourism destinations.

Collectively, these contributions demonstrate the broader methodological, theoretical, and practical implications of this study. The methodology can be replicated across contexts, the ontology advances conceptual clarity for future research, and the scorecard provides a decision-support tool aligned with international standards for smart tourism governance.

## 3 Methodology

The methodology applied in this study is based on two widely recognized frameworks for data analysis: the cross-industry standard process for data mining (CRISP-DM) and knowledge discovery in databases (KDD). Both guarantee a rigorous, orderly, and effective process to transform data into useful knowledge for decision-making in the tourism sector (Camacho-Ruiz et al. 2023; de Jesús Carranza González et al. 2024; Huda et al. 2024).

The proposed methodology follows the CRISP-DM framework (Chapman et al. 2000) and integrates the key stages of the KDD framework (Fayyad et al. 1996) in an iterative manner (Fig. 2). The former (stages marked with squares) provides a structured and industry-proven framework for data mining projects, whereas the latter (stages marked with circles) focuses on the overall process of discovering knowledge from data.

Although CRISP-DM remains the de facto standard for developing data mining and knowledge discovery projects, the field has evolved significantly over the past two decades (Martínez-Plumed et al. 2020; Tripathi et al. 2021). This evolution has led to the need for more flexible models, particularly for exploratory data science projects.

The integration of CRISP-DM and KDD can help address this need by incorporating the structured approach of CRISP-DM with the knowledge discovery focus of KDD and providing a more adaptable framework that can accommodate both goal-directed and exploratory approaches to data analysis. Their integration allows for a more comprehensive and flexible approach to data analysis, addressing both the practical implementation aspects of CRISP-DM and the knowledge discovery goals of KDD.

Hence, by combining the strengths of both approaches, data scientists and analysts can develop more effective and efficient strategies for extracting valuable insights from large and diverse datasets, offering several advantages for data mining and knowledge discovery projects (Plotnikova et al. 2020). Ultimately, this innovative dual methodological approach (CRISP-DM and KDD) ensures the rigorous, valid, and efficient analysis of tourism sector data supported by proven and accepted practices.

### 3.1 Business understanding

Having identified the need to measure the online reputation of tourist destinations by encompassing all attraction points in a given geographical area. To achieve this goal, we developed a synthetic index with a quantitative, standardized, and human-centered approach.

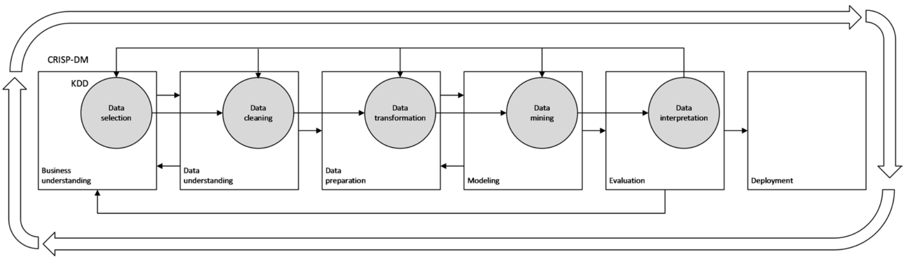


**Fig. 2** Stages of the methodological framework. Prepared by the authors

For the first stage of the KDD framework, Google Maps was selected as the primary data source. It is among the most widely used web mapping platforms and consumer applications, with more than 1 billion monthly active users worldwide and more than 25 million in Spain. This makes it a rich and accessible repository of real-time geographic and user-generated data.

Google Maps was chosen over other alternatives because of its popularity, its geolocated approach, and its free and collaborative nature. Users can share opinions and comments on each attraction point within a destination, which enriches the dataset and facilitates a comprehensive understanding of tourist perceptions (Ahmad et al. 2013; García Castillo et al. 2024; Mukherjee and Zalani 2024).

### 3.2 Data understanding

This stage involves the extraction, review, and preliminary evaluation of the quality and structure of the data collected from Google Maps, identifying the patterns, formats, and potential problems. Data extraction from Google Maps was carried out using the Apify Google Maps Scraper, which offers functionalities for accessing geographic and location data.

The process begins by collecting pins or markers available at the chosen destination. Each marker reflects an attraction point that has a uniform resource locator (URL) that is used for the final data extraction. All the data from the tourist reviews of each attraction point were extracted. An example of the data obtained from a tourist review of an attraction point is presented in Table 2.

As can be seen in the table, the data associated with a tourist review for a specific attraction point include both review-specific information (first five rows of Table 1) and attributes related to the attraction itself (remaining rows of Table 1). The TotalScore field represents the average rating of the attraction point, adding the rating of the reviews multiplied by the number of reviews of each review distribution (ReviewsDistribution) and dividing it by the total number of reviews (ReviewsCount) of the attraction point.

Finally, the data were cleaned to integrate the second stage of the KDD framework using the statistical and visualization programming language R v4.3.2. Its consistency is verified (e.g., homogeneous formats, elimination of duplicates, redundant, inconsistent, or incomplete data), and its quality is reviewed (e.g., detection of null values or outliers and correction of missing values). For this purpose, certain R packages are used that provide functions and tools for data processing, such as core Tidyverse to manipulate data and Readxl to read the data from the Microsoft Excel © file extracted from Google Maps.

In this way, to enhance the transparency, replicability, and interpretability of the study, Appendix I presents a brief summary of the extracted data for the municipality of Burgos, which will be used later for the case study in Sect. 4. As an example, only 10 reviews are shown, one for a different attraction point. A total of 476,701 reviews were obtained for 1,792 attraction points in that municipality.

**Table 2** Data extracted for a tourist review

| Name | Type | Data |
|---|---|---|
| Id_review | String | ChdDSUhNMG9nS0VJQ0FnSURWdHB1YnpRRRAB |
| Name_reviewer | String | Isa Martínez |
| Text_review | String | A good place to stop for breakfast and/or lunch |
| Stars | Integer (1–5) | 5 |
| PublishedAtDate | Datatime (yyyy-MM-dd HH: mm: ss) | 2023-12-12T22:58:56.790Z |
| Id_attraction_point | String | 440,644,646,303,454,194 |
| Name_attraction_point | String | B&J EuroCafe |
| Location/lat | Float | 42.3451845 |
| Location/lng | Float | -3.6978598 |
| ReviewsCount | Integer | 84 |
| ReviewsDistribution/fiveStar | Integer | 39 |
| ReviewsDistribution/fourStar | Integer | 28 |
| ReviewsDistribution/threeStar | Integer | 13 |
| ReviewsDistribution/twoStar | Integer | 0 |
| ReviewsDistribution/oneStar | Integer | 4 |
| TotalScore | Float (1–5) | 4.2 |
| WheelchairAccesibleEntrance | Boolean (True/False) | True |
| ChildFriendly | Boolean (True/False) | True |
| ClaimBusiness | Boolean (True/False) | True |

### 3.3 Data preparation

Using the same programming language, the data were then transformed to make them suitable for further analysis, thus integrating the third stage of the KDD framework. To construct the visitor profile, the user's sex (male and female) obtained from the user's name (Name_reviewer) was considered. The Gender package was used for this purpose. Notably, because it is a username, not everyone identifies themselves with a real name. The language in which the reviews were written (Spanish, English, French, Chinese, German, and others) was also considered (Text_review), which allowed the identification of the origin of the tourists who wrote the review. The Textcat package was used in this case. It is important to note at this point that the visitor profile is not used in the definition of TORI (Eq. 1); it simply allows the definition of the visitor profile in the bottom-left section of the constructed scorecard (Fig. 6). Consequently, since it is an interactive scorecard, all the metrics shown, including TORI, are recalculated considering only the values associated with the reviews of visitors with the selected profile, thus facilitating decision-making for each visitor segment. For example, if a male-English profile is selected, all the metrics will show the corresponding values when only the reviews of tourists who meet this profile are considered.

Defining an ontology to assign a certain category to each attraction point has also been considered. To achieve this, the categorization proposed by Defert (1980), which is based on Aristotle's cosmology, was used as a starting point. Defert defines four types of attraction points: (1) *Hydromo*, which encompasses all the elements related to water, whether in its natural state or modified by human intervention; (2) *Phytomo*, which refers to terrestrial elements and covers everything that has a certain natural attractiveness, whether modified by human activity or not; (3) *Lithomo*, which includes all the elements constructed by humans that are of interest either because of their nature or the use for which they are intended; and (4) *Anthropomo*, which focuses on the human being as a fundamental element, including everything related to socioeconomic structure, crafts, folklore, gastronomy, cultural activities, etc.

From this classification, a more detailed ontology was defined (Fig. 3), constituting categories, subcategories, and sub-subcategories, which allows each specific attraction point to be reached at this last level. Thus, this constitutes one of the outputs generated by the methodology.

Considering the municipality of Burgos, Fig. 4 shows the volume and distribution of the attraction points on the basis of the types defined by Defert (top left) and the categories (top right), subcategories (bottom left), and sub-subcategories (bottom right) proposed in this study. As can be seen, most of the attraction points belong to the sub-subcategories of Goods and Restoration.

It is also worth highlighting that in terms of distribution by category, the nature category encompasses the *Hydromo* and *Phytomo* attraction points, which is why a neutral color is used. Similarly, in the distribution by sub-subcategories, the sum of the attraction points that have not been assigned to any sub-subcategory, i.e., the *Lithomo* attraction points, appears in a neutral color.

With respect to another of the outputs generated by the methodology, a textual analysis of tourist reviews was undertaken to extract insights into their experiences at the respective attraction points. Thus, NLP techniques were applied to conduct sentiment analyses of the tourists. The specific package for text analysis, Tidytext, was used for this purpose. Specifically, the AFINN lexicon was used, which has 2,476 words in Spanish with a numerical rating of -4 to 4, depending on their polarity (more negative or more positive).

After the text was tokenized, the polarity function was calculated for each review; that is, on the basis of the lexicon, the ratings of the words were added together, resulting in a final rating. As a result, reviews are classified as positive or negative on the basis of the polarity function. If it is greater than 0, it is considered positive, whereas if it is less than 0, it is considered negative.

As an example, let us consider the text from the Text_review field (Table 2). The Tidytext package allows tokenizing the review, that is, separating the words and removing the stop words. Each word is subsequently scored using the dictionary provided by the AFINN lexicon (reflecting a $word \rightarrow score$ match). In this case, only one of the 9 words in the review is scored by the dictionary, specifically the word good, whose score es$+3$ ($good \rightarrow +3$). Finally, polarization is calculated as the

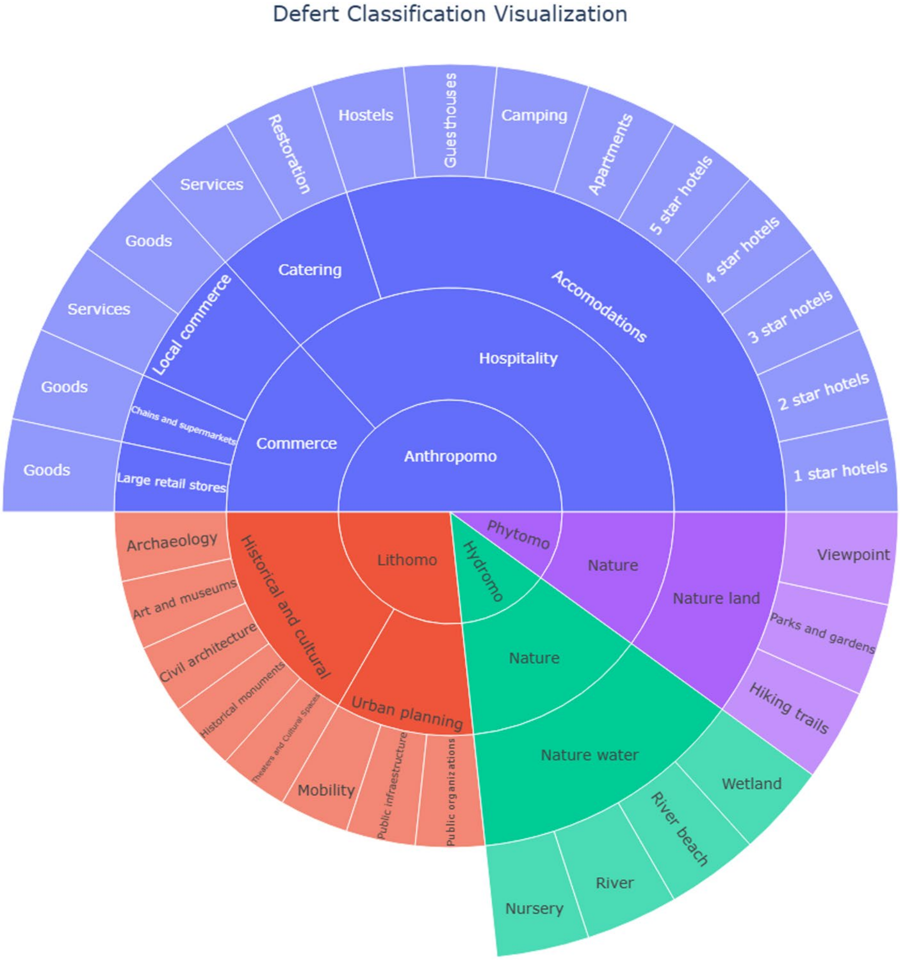


**Fig. 3** Ontology for attraction points. Prepared by the authors

sum of all the scores of the scored words divided by the number of the scored words, i.e., $3/1$. As a result, the polarization of this review is $+3$, indicating clearly positive sentiment. The polarization function subsequently allows this quantitative value of tourist sentiment to be integrated both into the definition of TORI (Eq. 1), which is presented in the top-left section of the constructed scorecard (Fig. 6), and into the bottom-right section of the scorecard specifically intended for sentiment analysis.

### 3.4 Modeling

The technical process of data mining, which makes sense of the data considered thus far and converts it into information and subsequently into knowledge, is divided into two subsections: data modeling and the definition of the synthetic tourism online reputation index. Thus, the fourth stage of the KDD framework is integrated.

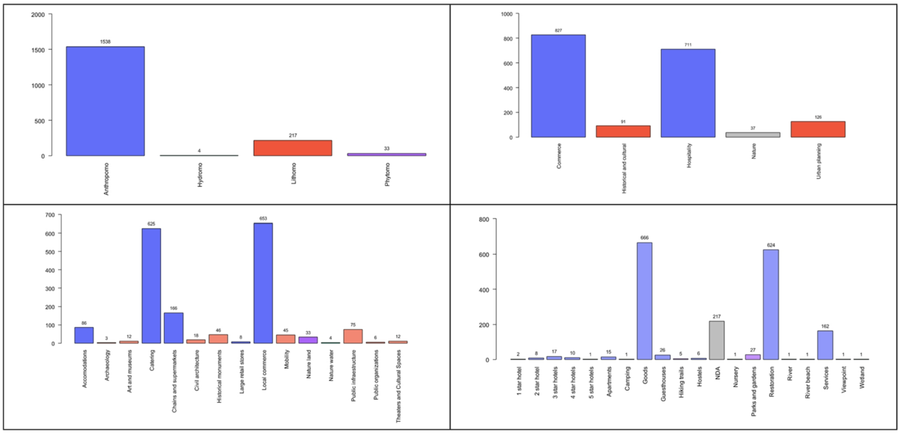


**Fig. 4** Volume and distribution of attraction points by types, categories, subcategories, and sub-subcategories. Prepared by the authors

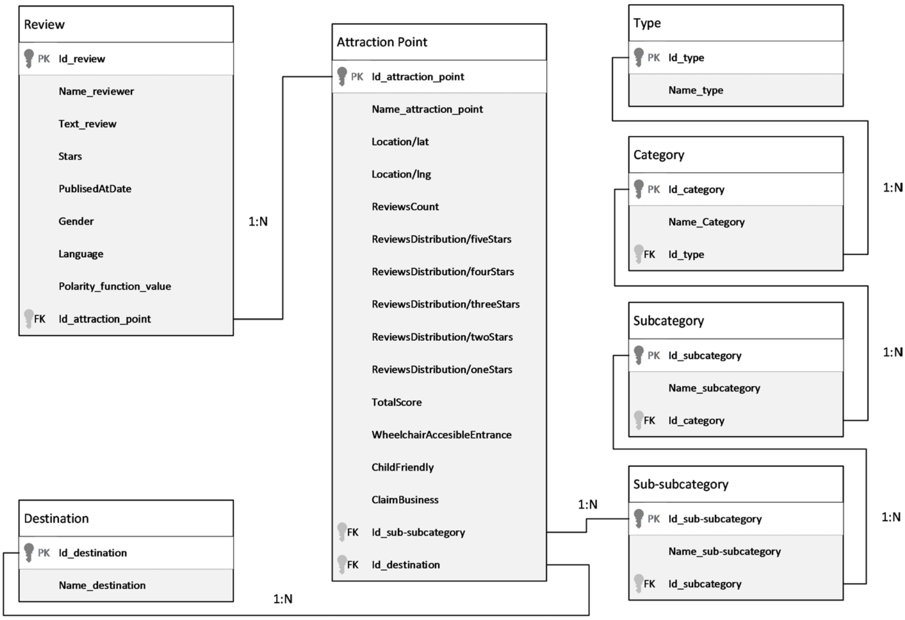


**Fig. 5** Relational data model. Prepared by the authors

### 3.4.1 Data model

A relational data model was defined (Fig. 5) using the data extracted from Google Maps (Table 2) and the defined ontology (Fig. 3). The Review table stores data related to each review of an attraction point. Each attraction point can include multiple reviews, while each is located at a single tourist destination. Moreover, one destination can encompass several attraction points.

In terms of classification, each attraction point belongs to a sub-subcategory. In turn, each sub-subcategory can contain multiple attraction points. The same hierarchical relationship applies to subcategories, categories, and types, following Defert categorization.

Note that the Review table displays the calculated fields, Gender, Language, and Polarity_function_value, which contain the final rating of the polarity function. The remaining fields contained primary data.

### 3.4.2 Tourism online reputation index (TORI)

Defining a synthetic index to measure the online reputation of a tourist destination constitutes a key component in fulfilling the main objective of this study. To calculate the tourism online reputation index (TORI), the TotalScore and the average of the polarization function of a tourist destination's attraction point, number of reviews, number of attraction points, and their categorization into types, categories, subcategories, and sub-subcategories are used.

The TORI aims to guarantee a more representative measure of overall tourist satisfaction by weighting the TotalScore plus the average of the polarization function of the attraction point on the basis of the number of reviews it has and considering the total number of reviews for either its destination, its categorization, or both. This approach allows attraction points with more reviews to have a greater influence on the index, reflecting the opinions of a greater number of tourists.

In contrast, attraction points with fewer reviews have a proportionally smaller influence on the index calculation. In addition, to achieve greater granularity, a scale of 1 to 10 is used instead of the scale offered by Google Maps of 1 to 5 (Stars), adding to the TotalScore value of each attraction point the average of the polarization function normalized to 5 (in this case, since the range of values of the AFINN lexicon is $[-4, +4]$, the denominator of the normalization is 4. If it were $[-5, +5]$, then the denominator would be 5). This increasing granularity provides a broader view of tourist satisfaction. The mathematical notation used for its calculation is explained below (the category is used as a categorization of the attraction point, but the types, subcategories, and sub-subcategories are similar):

Let:

$D$: Set of destinations. $D = \{d_1, d_2, \dots, d_n\}$, where $n$ denotes the number of destinations.

$C$: Set of categories. $C = \{c_1, c_2, \dots, c_m\}$, where $m$ denotes the number of categories.

$AP_{d_n c_m}$: Subset of attraction points from destination $d_n$ and category $c_m$.

$AP$: Set of attraction points. $AP = \{AP_{d_1 c_1}, AP_{d_1 c_2}, \dots, AP_{d_n c_m}\}$.

$ap_{inm}$: Specific attraction point in $AP_{d_n c_m}$.

$TS\,(ap_{inm})$: TotalScore of the attraction point $ap_{inm}$.

$PF\,(ap_{inm})$: Average value of the polarity function of the attraction point $ap_{inm}$ reviews.

$N\,(ap_{inm})$: Number of reviews (ReviewsCount) of attraction point $ap_{inm}$.

Thus, the TORI of a given tourist destination and a given category $TORI\left(d_n, c_m\right)$ is defined in Eq. (1). If instead of a single tourist destination and a single category, it is necessary to analyze several tourist destinations and/or several categories, and $d_n$ and/or $c_m$ must be updated.

$$TORI\left(d_n, c_m\right) = \frac{\sum_{ap_{inm} \in AP_{d_n, c_m}} \left(TS\left(ap_{inm}\right) + \frac{5}{4} * PF\left(ap_{inm}\right)\right) * N\left(ap_{inm}\right)}{\sum_{ap_{inm} \in AP_{d_n, c_m}} N\left(ap_{inm}\right)} \quad (1)$$

Thus, the TotalScore plus the average of the polarization function of each attraction point is weighted by its relevance, represented by its number of reviews, is increased in granularity, and is normalized by the total number of reviews for its destination and category. This means that regardless of the number of attraction points for each destination and category or how many reviews are accumulated in total, the index is scaled proportionally.

For example, if a destination and category have many reviews, the denominator will be larger, reducing the relative impact of each individual attraction point in the index. Conversely, if a destination and category have few reviews, the denominator will be smaller, thereby increasing the relative influence of each attraction point in the index. Therefore, the denominator ensures that the index is proportional to the relative weight of each attraction point in each destination and category on the basis of the number of reviews. Therefore, in general terms, the TORI for all destinations and all categories is defined in Eq. (2).

$$TORI = \frac{\sum_{ap_{inm} \in AP} \left(TS\left(ap_{inm}\right) + \frac{5}{4} * PF\left(ap_{inm}\right)\right) * N\left(ap_{inm}\right)}{\sum_{ap_{inm} \in AP} N\left(ap_{inm}\right)} \quad (2)$$

This index weighs the TotalScore plus the average of the polarization function of the attraction points on the basis of their relative importance (based on the number of reviews) within the entire set of destinations and within the entire set of categories. Finally, the TORI of a single attraction point is defined in Eq. (3).

$$TORI\ \left(ap_{inm}\right) = TS\left(ap_{inm}\right) + \frac{5}{4} * PF\left(ap_{inm}\right) \quad (3)$$

In this case, since the number of reviews of the attraction point $N\left(ap_{inm}\right)$ coincides with the number of reviews of the total number of selected attraction points (denominator of Eq. (1) and/or (2)), both factors can be eliminated so that the TORI is obtained by simply adding the TotalScore of the attraction point with its average of the polarity function normalized to 5.

### 3.5 Evaluation

To facilitate data interpretation (integration of the fifth state of the KDD framework), the methodology is validated and critically reviewed by applying a specific case study (Sect. 4), thus ensuring the representativeness and accuracy of the generated synthetic TORI. The index is interpreted according to the well-known net promoter score (NPS) (Reichheld 2003), a widely used market research metric for online reputation management that is used as an indicator of the overall perception of tourist satisfaction (Kushcheva and Eilola 2023; Antón-Maraña et al. 2023).

Ratings less than or equal to 7 are related to very low levels of satisfaction and perception and, therefore, a priority area for reconfiguration to improve the reputation of the destination and prevent negative recommendations. Ratings greater than 7 and less than or equal to 9 were associated with good levels of satisfaction and perception, although this is an area with potential improvement. The reputation of the destination may be affected if the level of trust is not maintained. Finally, ratings greater than 9 and less than or equal to 10 indicated excellent levels of satisfaction and perception.

In some cases, there is room for improvement, although in most cases, these levels need to be maintained and consolidated. The reputation of a destination reflects positive recommendations from tourists, thus demonstrating loyalty. Furthermore, quantitative (supply and demand indicators) and qualitative (sentiment analysis) results were analyzed to comprehensively reflect perceptions of tourism.

### 3.6 Deployment

Finally, an interactive visualization is designed in a scorecard using the Microsoft Power BI © tool (to construct calculation formula expressions, the library of functions and operators data analysis expressions (DAX) has been used), facilitating the interpretation of the results and their practical use by decision-makers in the tourism industry. The scorecard stands as the last of the outputs generated by the methodology.

## 4 Results: case study

The province of Burgos (Spain) has been identified as a target area owing to its dynamic tourism sector and the challenges it faces in terms of depopulation and sustainable tourism development. With a population density of 25.239 inhabitants/km$^2$, Burgos is a sparsely populated territory (between 12.5 and 50 inhabitants/km$^2$) according to the European Union Standards (Europarl 2025; Eurostat 2025). Population decline not only has demographic effects but also negatively impacts social, cultural, environmental, and economic aspects, weakening the social fabric and reducing communities' ability to sustain and develop (Serra et al. 2022). In this regard, the tourism sector has established itself as a key factor in the economic and social development of the region, with an impact of 414.3 million euros and the creation of 11,950 jobs by 2024 (JCyL 2025), and as a possible solution to mitigate the problem of depopulation, at least partially.

Thus, strengthening the technological axis defined by the SEGITTUR model (see Sect. 1) can facilitate the acquisition of the STD distinction by the province of Burgos (currently, it does not reach 80% compliance with the established requirements). Only seven municipalities in Spain achieve this distinction, all coastal), which is expected to have a positive effect on the regional economy and a direct effect on the depopulation problem, especially in inland areas such as this one.


**Fig. 6** Scorecard for the province of Burgos. Prepared by the authors

Different studies have demonstrated the impact of tourism on the problem of depopulation (Cáceres-Feria et al. 2021; Hashimoto et al. 2021; Vidal-Matzanke and Vidal-González 2022), and more generally, studies have also analyzed the importance of the tourism sector in inland areas (Jesus and Franco 2016; Scorza et al. 2019; Prat Forga 2020; Baptista Alves et al. 2022; Gatto et al. 2022), demonstrating that, currently, tourism has become a key activity that has a growing influence on both the economy and the society of the territories (Urrutia et al. 2018).

Following the methodology designed throughout this work, this section presents the TORI of the province of Burgos, together with related information on supply, demand, and sentiment analysis (Fig. 6). Thus, in addition to reinforcing the technological axis, this methodology is validated, and its applicability in a real environment is evaluated, allowing verification of its effectiveness and capacity to provide relevant information that can be used in strategic decision-making in the tourism sector.

The scorecard for the province of Burgos initially displays, in the top-left section, information related to the tourism online reputation index (TORI). It allows users to select the destination, either one or several of the 371 municipalities comprising the province as well as the category, subcategory, and sub-subcategory for analysis. The municipality of Burgos is selected for this case study. Moreover, if a category is selected, the information is disaggregated into subcategories and so forth. The TORI value (8.42), number of attraction points (1,792), and number of reviews (476,701) are calculated on the basis of the four previous selections. It also shows the TORI value broken down by category and the evolution of the TORI since 2019 (annual and cumulative).

A TORI score of 8.42 out of 10 indicates that the evaluated tourist destination has a good level of satisfaction and perception. However, there is still room for improvement to consolidate its reputation and avoid potential negative impacts in the future.

The *Historical and cultural* category obtained the best score (9.34), placing it in the range of excellence. Nevertheless, if this information is combined with the percentage of reviews by category (bottom-left section), which shows that this category represents 13.42% of the total reviews, it can be deduced that tourists highly value this aspect but that its positive contribution to the overall TORI is limited. Actions should be taken to maintain and consolidate this rating, as it reflects strong tourist loyalty and satisfaction.

With respect to the *Nature* category, the TORI is on the verge of excellence (8.92), with a very high score, although it only accounts for 2.26% of the total reviews. As in the previous case, although its quantitative impact is limited, it is important to maintain and slightly improve the attractiveness of this category to achieve high levels of excellence.

The *Hospitality* and *Commerce* categories obtained good scores, 8.24 and 8.34, respectively. Furthermore, these categories are particularly relevant because they account for 61.05% and 20.95% of the total volume of reviews, respectively. This finding indicates that although the current perception is positive, the destination has significant room for improvement in these categories, given that any deterioration could significantly affect its online reputation, especially in the *Hospitality* category, owing to its dominant quantitative weight.

Finally, the *Urban planning* category presents low representativeness (2.33%) and the lowest score (8.02); thus, it would be advisable to assess the specific aspects that influence this perception to consider possible specific actions that could indirectly benefit the overall TORI of the tourist destination.

In the top-right section of Fig. 6, the metrics related to the tourist offer (i.e., the attraction points) can be seen. The number of attraction points in the initial selection (top left panel) and the percentage of attraction points in the selection (categories, subcategories, and sub-subcategories) compared to the total number of attraction points in the selected municipalities are shown. The number of attraction points for which information has been requested and updated by owners or managers is also shown. A total of 70.48% of the attraction points were requested, indicating a certain degree of reliability in the information on tourist offerings.

Below is the distribution by category of attraction points (46.15% *Commerce*, 39.68% *Hospitality*, 7.03% *Urban planning*, 5.08% *Historical and cultural*, and 2.06% *Nature*) and their geolocation (the map shows the number of reviews). Finally, two metrics are used to measure the destination's inclusiveness: child-friendliness (13.54% of attraction points are ideal for family visits with children) and wheelchair accessibility (45.07% of attraction points are adapted). It should be noted that a high percentage of attraction points do not provide information on these metrics (86.46% and 53.59%, respectively); therefore, action is needed.

The bottom part of Fig. 6 is related to tourist demand. The left-hand side shows the number of reviews. This metric reflects the interest in the attraction points of a tourist destination. The greater the number of reviews, the greater the interest in conveying greater trust and credibility. The tourist profile is represented by the language used (93.18% Spanish speaking) and sex (54.12% men). The number of reviews by category and the evolution of the average ratings (Stars) by year are shown next.

Below is a series of metrics that break down the number of reviews by year (showing their evolution), month and year, and day. Finally, the distribution of reviews and the distribution of reviews by category based on ratings are also presented. It is worth noting that the attraction points in the *Historical and cultural* category are the highest rated (74.34% of reviews are five stars), whereas those in the *Urban planning* category are the lowest (25.66% of reviews are one, two, or three stars). Clearly, actions must be taken in this category.

Finally, on the right-hand side, a sentiment analysis can be found. An analysis of positive or negative sentiments can be performed separately. Next, the average rating (3.36), the number of reviews that contain text, and the number of reviews with text by category are shown.

Below, the distribution of the polarity function values shows the percentage of positive (90.53%) and negative (9.47%) reviews. Finally, the most frequently used words (word cloud) are also shown along with the sentiment analysis by category. In the latter, the greatest number of reviews with positive sentiment comes from the *Hospitality* category, although it is also true that this same category generates the most negative sentiment.

All these results can be validated, at least partially, by comparing them with the degree of tourist satisfaction offered by Antón Maraña et al. (2021) and in more detail by the Burgos Tourism Observatory (BTO, 2019). This study uses 1,554 observations obtained through a survey conducted at different points in the province of Burgos. The survey points were selected on the basis of the touristic potential index (TPI), which allows the identification of destinations with greater tourist attraction (Aparicio Castillo et al. 2023). The study indicated a degree of overall satisfaction of 4.57 out of 5 points, which can be converted to 9.14 out of 10 points, a value similar to the 8.42 offered by TORI. Similarly, in the categories of *Historical and cultural* (8.52 vs. 9.34), *Hospitality* (8.30 vs. 8.24), and *Commerce* (7.46 vs. 8.34), similar values were obtained (survey vs. TORI). In the other categories, there was not enough data for comparison. It should be noted that this comparison should be interpreted with caution since the survey analyzes the degree of tourist satisfaction on a global trip. Furthermore, it should also be noted that the characteristics of the sample are different. For example, the survey was conducted directly on the street, while the reviews were obtained from Google Maps, which can generate a certain degree of anonymity and less involvement in the scores given by tourists.

Therefore, interpreting all these metrics, both separately and together by decision-makers, provides a broader view of the online reputation of the tourism sector in the city of Burgos. This also facilitates the identification of strengths, weaknesses, threats, and opportunities as well as a better understanding of tourist perceptions.

## 5 Discussion

The application of the TORI index to the municipality of Burgos demonstrates its capacity to address the key objective of this study: to develop a standardized, scalable and human-centered tool to evaluate the online reputation of tourist destinations. As anticipated in the conceptual framework, the integration of geolocated ratings, the number of reviews, and sentiment analysis allows for a multidimensional measurement that reflects both the intensity and the quality of visitor experiences across different types of attraction points. This aligns with the STD paradigm, which emphasizes data-driven governance and participatory evaluation mechanisms.

The empirical results also reinforce the conclusions of the literature review regarding the limitations of existing indices—namely, their fragmentation, lack of standardization, and narrow spatial scope (Bui et al. 2021; Carrasco-Santos et al. 2021; Pártlová et al. 2022; Qin et al. 2022; Antón-Maraña et al. 2023. TORI addresses these weaknesses by aggregating diverse types of user-generated content into a single synthetic metric.

In this way, the index goes beyond the evaluation of isolated attraction points and reflects the collective reputation of an entire destination. By doing so, it operationalizes theoretical insights into cocreation, digital engagement, and territorial inclusivity, contributing to the technological axis of the STD model and advancing methodological innovation in the field.

Moreover, the results directly support other outputs generated by the methodology. First, the ontology used to classify attraction points across the municipality proved essential for disaggregating and interpreting the reputation of various categories—such as cultural, natural, commerce, urban planning, hospitality-based attractions—as well as subcategories and sub-subcategories. This categorization facilitates more granular analysis and offers valuable insights for destination managers seeking to balance and promote their tourism assets.

Similar conclusions were drawn by Faerber et al. (2021), who highlighted that different attraction types—such as cultural, natural, and event-based—reveal distinct patterns in visitors' perceptions and experiences, underscoring the strategic value of fine-grained categorization for destination management and marketing decisions.

Second, the application of NLP to perform sentiment analysis added depth to the numerical ratings by revealing the emotional tone and qualitative perception of tourists. The sentiment-based dimension of the index enriches the understanding of destination image and aligns with the growing emphasis in the literature on incorporating subjective, experiential data into tourism analytics. Similarly, Álvarez-Carmona et al. (2022) found that integrating polarity measures into tourism analytics provides deeper insight into perceived authenticity and service quality.

Finally, the development of a dynamic scorecard offers a practical interface to visualize reputation indicators and monitor changes over time. This tool responds to calls in the literature for actionable metrics and supports informed, strategic decision-making by local stakeholders. It also strengthens the replicability and adaptability of the methodology to other destinations, thereby broadening the scope and relevance of the study's contributions. Comparable efforts were reported by Cillo et al. (2021) in niche tourism destinations, where big data analytics systems enabled managers to monitor tourist perceptions in real time, identify operational issues, and allocate resources more efficiently to enhance competitiveness.

### 5.1 Methodological implications

The proposed methodology, which integrates CRISP-DM and KDD, applies advanced data mining techniques to extract and analyze large volumes of user-generated content. By combining geotagging, tourist ratings, and sentiment analysis, the approach makes it possible to identify specific attraction points within a destination and to define an ontology for their systematic categorization.

This structured framework facilitates the creation of a synthetic index that is not only quantitative and standardized but also human centered. As a result, it provides a comprehensive and nuanced representation of a destination's online reputation that goes beyond isolated indicators.

Finally, integrating these components into a user-friendly scorecard offers stakeholders a practical decision-making tool. The scorecard enables them to monitor performance, improve the tourist experience, and optimize marketing strategies through accessible and data-driven insights.

## 5.2 Theoretical implications

From a theoretical standpoint, this study contributes to advancing the academic understanding of online reputation in the context of tourism. The proposed index goes beyond existing metrics by providing a more accurate reflection of a destination's reputation at different levels—whether at the scale of the whole destination, a category, a subcategory, or even an individual attraction point. It achieves this by simultaneously considering both ratings (scores plus polarity) and the total number of reviews linked to each unit of analysis.

A key theoretical contribution lies in how the index assigns greater weight to attraction points with higher volumes of reviews. This design improves the reliability of the overall reputation for two main reasons. First, a larger number of reviews increases statistical confidence: an attraction point with many reviews reflects a broader and more representative perception of tourists. Second, it mitigates the problem of scarce data: if an attraction point has few reviews, even with a low rating, its impact on the overall reputation remains limited and avoids distorting the aggregate assessment.

In this way, the index ensures that less relevant or nonrepresentative data are not overinterpreted. For example, a poorly rated attraction point with only a handful of reviews does not disproportionately damage the reputation of a destination that otherwise receives strong evaluations. This nuanced approach provides a more robust and fair representation of online reputation, aligning theoretical models with the realities of tourist-generated digital content.

Another theoretical contribution of the index is that it explicitly recognizes the dual importance of quality (ratings) and quantity (number of reviews). An attraction point with a high rating but very few reviews cannot strongly influence the overall reputation, as its reliability is limited (high quality–low quantity). Conversely, a site with many reviews but low ratings also fails to provide a balanced perception (low quality–high quantity). The index shows that moderate cases—where both the ratings and the volume of reviews are sufficient—tend to provide the most reliable representation of tourist perceptions. This highlights the theoretical need to jointly consider service quality and tourist participation to develop a robust reputation.

Moreover, the results indicate that combining quantitative data (e.g., number of reviews, average rating) with qualitative data (e.g., sentiment polarity from comments) produces a richer and more trustworthy measure of the tourist experience. This methodological integration strengthens theoretical models by linking objective indicators with subjective perceptions.

From an external perspective, the index also contributes to theory by reinforcing confidence in aggregated online reputation. A larger number of reviews signals greater consensus and minimizes bias from anecdotal or extreme evaluations. For example, a single negative review with limited visibility does not significantly alter the perceived reputation of a destination with hundreds of positive and consistent evaluations. Thus, it avoids misunderstandings about the overall reputation.

Finally, the study supports the view that reputation is a composite, place-based construct that emerges from the aggregation of digital interactions rather than isolated metrics. In this sense, the introduction of an ontology to classify attraction points adds theoretical depth, offering a structured approach to destination-level evaluation that has often been absent in the existing literature.

### 5.3 Practical implications

From a managerial perspective, the index and its accompanying scorecard provide clear priorities for action. For example, if an attraction point accumulates many reviews but receives a low rating, it requires urgent attention given its significant impact on overall reputation. Conversely, the methodology prevents sites with few or irrelevant reviews from disproportionately influencing the destination's reputation, allowing managers to focus resources on what truly matters.

A second implication concerns the interactive scorecard developed. This tool facilitates real-time monitoring and visualization of key indicators, which can be especially valuable for local governments, destination management organizations (DMOs), and tourism professionals. In doing so, it strengthens the technological and governance dimensions of the STD framework and supports evidence-based decision-making, particularly in rural and less developed areas.

The case study of Burgos illustrates how the methodology can inform practical strategies for destination managers and public institutions. Historical and cultural attractions as well as natural attractions receive highly positive evaluations but remain relatively invisible in terms of review volume. This suggests the need to enhance promotional strategies to increase the visibility of these high values but underexposed resources, especially in rural areas where tourism can foster inclusive and sustainable development. In contrast, hospitality and commerce attractions dominate in volume but show only moderate performance, highlighting the need for improvements in service quality, staff training, and reputation management. Urban infrastructure, the weakest category in terms of perception, highlights the importance of investing in urban design, accessibility, and public-space visitor experience.

Another implication relates to inclusivity and transparency. This study identified a gap in the availability of information about accessibility and family friendliness at many attraction points. Destination managers should encourage providers to update their digital profiles and include more detailed information. Doing so would improve tourist decision-making and align with the technological and social axes of the STD distinction promoted by SEGITTUR.

In summary, the TORI is a robust tool because it balances the quantity and quality of data to produce a trustworthy representation of reputation. It enhances the accuracy and reduces the bias of external perceptions by prioritizing representative attractions and mitigating the influence of outliers or sparse data. In practice, this tool fosters trust among tourists, support informed decision-making by managers, and contribute to achieving the STD distinction. In the case of Burgos, it contributes to boosting the technological axis of the STD model. Beyond strengthening competitiveness, it may also generate positive economic and social impacts in sparsely populated areas by promoting tourism as a driver of inclusive regional development.

## 6 Conclusions

The proliferation of social media platforms and review websites has amplified the impact of online reputation on tourism destinations. Destination management organizations and tourism businesses increasingly recognize the need to monitor and manage their online presence to maintain a positive image. This evolution has stimulated the development of sophisticated tools, strategies, and methodologies for analyzing sentiment, responding to feedback, and leveraging positive reviews to attract more tourists.

Against this backdrop, this study has achieved its main objective by developing a systematic and innovative methodology for constructing a synthetic tourism online reputation index (TORI). The proposed framework integrates big data techniques with geolocated information to capture tourists' perceptions of destination experiences through online ratings, reviews, and sentiment expressed across digital channels. Moreover, the results obtained with the TORI have been validated, at least partially, by comparing them with ground-truth data, thereby reinforcing the robustness and reliability of the index.

In fulfilling this objective, this research has produced several complementary outputs that enhance the methodological soundness of the approach. First, the development of an ontology enables a structured categorization of the tourism supply across attraction points, providing conceptual clarity and facilitating comparative analysis. Second, the application of NLP techniques supports sentiment analysis and tourist profiling, offering deeper insights into experiential dimensions that traditional indicators often overlook. Third, the design of an interactive scorecard allows for the visualization of key indicators, particularly the index itself, thereby strengthening the capacity of tourism stakeholders to incorporate data-driven evidence into strategic decision-making.

Taken together, these contributions offer a scalable and replicable approach for assessing online reputation at the destination level. By grounding the analysis in the full set of attraction points and leveraging large-scale digital information, the study advances current methodological practices and provides a robust tool for understanding and managing tourism reputation in increasingly data-rich environments. In doing so, it also addresses the main gaps identified in the literature: the need for governance models aligned with the STD paradigm and supported by real-time digital interaction; the limited integration of big data, NLP techniques, and geolocated information into comprehensive analytical frameworks; and the scarcity of destination-level approaches capable of quantifying online reputation through tourist-generated ratings, reviews, and sentiments. In response to these shortcomings, this study contributes to strengthening the conceptual and methodological foundations of reputation analysis in tourism research.

### 6.1 Work limitations and future directions

The limitations of this study are related mainly to data transformation (Sect. 3.3). Categorizing attraction points to fit the defined ontology is particularly difficult. Future research is needed to mitigate these limitations. For example, delving into NLP techniques, or other machine learning techniques, can facilitate the automatic categorization of attraction points by analyzing the names of attraction points and reviews provided by tourists. This can also help improve the proposed ontology.

Furthermore, in favor of a more human-centered approach, these techniques can help to identify or reduce the impact of reviews generated by robots or fake feedback, which could artificially increase the importance of certain attraction points and distort the results. Another limitation of the index is the problem that can be generated by a high number of reviews given to a particular attraction point; thus, its influence on the TORI can become disproportionately large, potentially overshadowing other significant attraction points but with few reviews.

Future research could define the maximum values for the weight assigned to an attraction point, although this would require decision-making by the index designer, limiting the influence of opinions expressed by tourists. Similar decision-making can be reviewed in Antón-Maraña et al. (2024) for the design of a synthetic index for the evaluation of the impact of events.

On the other hand, the inclusion of prediction engines and greater validation across diverse destinations could make the TORI a valuable tool for destination management organizations and policymakers, offering a data-driven approach to understanding and improving destination reputation. As the tourism industry continues to evolve in the digital age, such methodologies are becoming increasingly crucial for maintaining competitiveness and ensuring sustainable tourism development.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s40558-026-00368-0.

**Acknowledgements** This work was financially supported by the Spanish State Research Agency (Ministry of Science and Innovation of the Government of Spain) and FEDER Funds (Knowledge Generation Projects) through the project “Methodologies for solving problems with economic, social and environmental criteria. Application to healthcare resource management” (ECOSOEN-HEALTH, grant ref. PID2022-139543OB-C44). The authors sincerely appreciate the financial support received.Also, this work was financially supported by the Spanish Universities’ Ministry under the University Teachers’ Training Programme (FPU/00606-2021).

**Author contributions** J.C.P.R. and I.M.L. wrote the Introduction and context. J.C.P.R. contributed the Background and theoretical framework. J.C.P.R. and P.A.M. developed the Methodology. I.M.L. and P.A.V. wrote the Results: Case Study section. P.A.M. and P.A.V. contributed to the Conclusions, limitations, and future directions. J.C.P.R. and P.A.M. was responsible for the preparation of figures, tables, and English language revision. All authors reviewed and approved the final manuscript.

**Funding** Open access funding provided by FEDER European Funds and the Junta de Castilla y León under the Research and Innovation Strategy for Smart Specialization (RIS3) of Castilla y León 2021-2027.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

Absari DT, Djunaidy A, Susanto TD (2024) Development of evaluation readiness model for smart tourism destination. In: 2024 2nd International Conference on Software Engineering and Information Technology (ICoSEIT). IEEE, Bandung pp 311–317

Ahmad W, Zia A, Khalid U (2013) A Google Map Based Social Network (GMBSN) for Exploring Information about a Specific Territory. J Softw Eng App 6(7):343–348. https://doi.org/10.4236/jsea.2013.67043

Aïdi N, Fabry N (2024) Beyond the certification of smart tourism destination: insights from the city of Medellín in Colombia. Int J Tourism Cities 10(2):577–603. https://doi.org/10.1108/IJTC-03-2022-0056

Alhajri FN (2024) Current trends in natural language processing application and its applications in improving the quality of tourism communication. Int J Qual Res 18(3):807–816. https://doi.org/10.24874/IJQR18.03-11

Aliyah CL, Pangilinan GA, Chakim MHR, Saputra DB (2023) Examining the Impact of Artificial Intelligence and Internet of Things on Smart Tourism Destinations: A Comprehensive Study. Aptisi Trans Technopreneurship 5(2sp):12–22. https://doi.org/10.34306/att.v5i2sp.332

Alsharif A, Isa SM, Alqudah MN (2024) Smart Tourism, Hospitality, and Destination: A Systematic Review and Future Directions. J Tourism Serv 15(29):72–110. https://doi.org/10.29036/jots.v15i29.746

Álvarez-Carmona MÁ, Aranda R, Rodríguez-Gonzalez AY, Fajardo-Delgado D, Sánchez MG, Pérez-Espinosa H, Martínez-Miranda J, Guerrero-Rodríguez R, Bustio-Martínez L, Díaz-Pacheco Á (2022) Natural language processing applied to tourism research: A systematic review and future research directions. J King Saud Univ – Comput Inf Sci 34(10 B):10125–10144. https://doi.org/10.1016/j.jksuci.2022.10.010

Antón Maraña P, Aparicio Castillo S, Puche Regaliza JC, Arranz Val P (2021) A comparative study for understanding the impact and evolution of the tourism role in Burgos province. J Tourism Dev 36(1):221–231. https://doi.org/10.34624/rtd.v1i36.8697

Antón-Maraña P, Baruque-Zanón B, Porras-Alfonso S, Arranz-Val P (2023) Tourism Reputation Index for Assessing Perceptions on Destinations Using Collaborative Text Data. Stud Appl Econ 41(1). https://doi.org/10.25115/sae.v41i1.9076

Antón-Maraña P, Díez-Hernández J, Puche-Regaliza JC (2024) A Novel Multiperspective and Multistakeholder Approach to Event Impact Measurement: Synthesizing the Triple Bottom Line. Event Manag 28(6):819–841. https://doi.org/10.3727/152599524X17108153015691

Aparicio Castillo S, Arranz Val P, Alvear González A, Antón Maraña P, Puche Regaliza JC (2023) Target points as locations for conducting surveys: Touristic Potential Index of the province of Burgos. Rev Métodos Cuant Econ Empresa 36:1–16. https://doi.org/10.46661/rev.metodoscuant.econ.empresa.7798

Aransyah MF, Hermanto B, Muftiadi A, Oktadiana H (2025) Exploring sustainability oriented innovations in tourism: insights from ecological modernization, diffusion of innovations, and the triple bottom line. Cogent Soc Sci 11(1):2447396. https://doi.org/10.1080/23311886.2024.2447396

Baggio R, Cooper C (2010) Knowledge transfer in a tourism destination: the effects of a network structure. Servic Ind J 30(10):1757–1771. https://doi.org/10.1080/02642060903580649

Baggio R, Sainaghi R (2011) Complex and chaotic tourism systems: towards a quantitative approach. Int J Contemp Hosp Manag 23(6):840–861. https://doi.org/10.1108/09596111111153501

Ballina FJ (2022) Smart business: the element of delay in the future of smart tourism. J Tourrism Futur 8(1):37–54. https://doi.org/10.1108/JTF-02-2020-0018

Baptista Alves HM, Pires Manso JR, da Silva Serrasqueiro Teixeira ZM, Santos Estevão CM, Pinto Nave AC (2022) Tourism-based regional development: boosting and inhibiting factors. Anatolia: Int J Tourism Hosp Res 33(1):128–142. https://doi.org/10.1080/13032917.2021.1924211

Bastidas-Manzano A-B, Sánchez-Fernández J, Casado-Aranda L-A (2021) The Past, Present, and Future of Smart Tourism Destinations: A Bibliometric Analysis. J Hosp Tourism Res 45(3):529–552. https://doi.org/10.1177/1096348020967062

Berné-Manero C, García-González M, García-Uceda ME, Múgica-Grijalba JM (2011) La influencia de las TIC en la estructura del sistema de distribución turístico. Cuad Tur 28:9–22

Borges-Tiago MT, Avelar S (2025) Co-creation dynamics in tourism and hospitality: a horizon 2050 paper. Tourism Rev 80(1):194–208. https://doi.org/10.1108/TR-06-2024-0500

Bosangit C, Mena M (2009) Meanings, motivations and behaviour of filipino-american first-time visitors of the Philippines: A content analysis of travel blogs. In: International Conference on Tourist Experiences, Lancaster

BTO (2019) Burgos Tourism Observatory Inform 2018. Burgos Tourism Observatory. https://www.ubu.es/sites/default/files/portal_page/files/informe_2018.pdf. Accessed 08 February 2026

Buhalis D, Amaranggana A (2015) Smart Tourism Destinations Enhancing Tourism Experience Through Personalisation of Services. In: Tussyadiah I, Inversini A (eds) Information and Communication Technologies in Tourism. Springer International Publishing, Cham, pp 377–389. https://doi.org/10.1007/978-3-319-14343-9_27

Buhalis D, Law R (2008) Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. Tourism Manag 29(4):609–623. https://doi.org/10.1016/j.tourman.2008.01.005

Buhalis D, Sinarta Y (2019) Real-time co-creation and nowness service: Lessons from tourism and hospitality. J Travel Tourism Mark 36(5):563–582. https://doi.org/10.1080/10548408.2019.1592059

Buhalis D, Harwood T, Bogicevic V, Viglia G, Beldona S, Hofacker C (2019) Technological disruptions in services: lessons from tourism and hospitality. J Serv Manag 30(4):484–506. https://doi.org/10.1108/JOSM-12-2018-0398

Buhalis D, Leung XY, Fan D, Darcy S, Chen G, Xu F, Wei-Han Tan G, Nunkoo R, Farmaki A (2023) Editorial: Tourism 2030 and the contribution to the sustainable development goals: the tourism review viewpoint. Tourism Rev 78(2):293–313. https://doi.org/10.1108/TR-04-2023-620

Bui V, Alaei AR, Vu HQ, Li G, Law R (2021) Revisiting tourism destination image: A holistic measurement framework using big data. J Travel Res 61(6):1287–1307. https://doi.org/10.1177/00472875211024749

Bulchand-Gidumal J (2020) Impact of Artificial Intelligence in Travel, Tourism, and Hospitality. In: Xiang Z, Fuchs M, Gretzel U, Höpken W (eds) Handbook of e-Tourism. Springer International Publishing, Cham, pp 1–19. https://doi.org/10.1007/978-3-030-05324-6_110-1

Cáceres-Feria R, Hernández-Ramírez M, Ruiz-Ballesteros E (2021) Depopulation, community-based tourism, and community resilience in southwest Spain. J Rural Stud 88:108–116. https://doi.org/10.1016/j.jrurstud.2021.10.008

Camacho-Ruiz M, Carrasco RA, Fernández-Avilés G, LaTorre A (2023) Tourism destination events classifier based on artificial intelligence techniques. Appl Soft Comput 148:110914. https://doi.org/10.1016/j.asoc.2023.110914

Carrasco-Santos MJ, Ciruela-Lorenzo AM, Méndez-Pavón JG, Cristófol-Rodríguez C (2021) An online reputation analysis of the tourism industry in Marbella: A preliminary study on open innovation. J Open Innov: Technol Mark Complex 7(2):111. https://doi.org/10.3390/joitmc7020111

Carvalho P, Alves H (2023) Customer value co-creation in the hospitality and tourism industry: a systematic literature review. Int J Contemp Hosp Manag 35(1):250–273. https://doi.org/10.1108/IJCHM-12-2021-1528

Cavalheiro MB, Joia LA, do Canto Cavalheiro GM (2020) Towards a smart tourism destination development model: promoting environmental, economic, socio-cultural and political values. Toursim Plan Dev 17(3):237–259. https://doi.org/10.1080/21568316.2019.1597763

Cavalheiro MB, Joia LA, do Canto Cavalheiro GM, Mayer VF (2021) Smart tourism destinations: (Mis) aligning touristic destinations and smart city initiatives. BAR-Bra Adm Rev 18(1):e190132. https://doi.org/10.1590/1807-7692bar2021190132

Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, Wirth R (2000) CRISP-DM 1.0: Step-by-step data mining guide. SPSS inc. https://mineracaodedados.files.wordpress.com/2012/12/crisp-dm-1-0.pdf. Accessed 05/08/2025

Chen S, Wang X, Zhang H (2021) Customer purchase forecasting for online tourism: A data-driven method with multiplex behavior data. Tourism Manag 87:104357. https://doi.org/10.1016/j.tourman.2021.104357

Cheng X, Xue T, Yang B, Ma B (2023) A digital transformation approach in hospitality and tourism research. Int J Contemp Hosp Manag 35(8):2944–2967. https://doi.org/10.1108/IJCHM-06-2022-0679

Chia SKS, Lo MC, Razak ZB, Wang YC, Mohamad AA (2021) Impact of destination image on tourist satisfaction: the moderating effect of Information Technology (IT). Geo J Tourism Geosites 34(1):88–93. https://doi.org/10.30892/gtg.34112-623

Cillo V, Rialti R, Giudice MD, Usai A (2021) Niche tourism destinations' online reputation management and competitiveness in big data era: evidence from three Italian cases. Curr Issues Tourism 24(2):177–191. https://doi.org/10.1080/13683500.2019.1608918

de Jesús Carranza González C, Moheno GA, de los Santos Torres G (2024) The Impact of Information Technologies on the Tourist Experience: Case of Tabasco, Mexico. In: Abreu A, Carvalho JV, Liberato P, Monroy HC (eds) Advances in Tourism, Technology and Systems. ICOTTS 2023. Smart Innovation, Systems and Technologies. Springer, Singapore, pp 283–292. https://doi.org/10.1007/978-981-99-9765-7_21

de López A (2015) Smart destinations: XXI century tourism. In: ENTER2015 conference on information and communication technologies in tourism. EAIT, Lugano

Defert P (1980) Essai systématique d'un inventaire des ressources touristique. Tourism Rev 35(1):14–15

Díaz E, Esteban Á, Koutra C, Almeida S, Carranza R (2023) Co-creation of value in smart ecosystems: past trends and future directions in tourism literature. J Hosp Tourism Technol 14(3):365–383. https://doi.org/10.1108/JHTT-04-2021-0122

Díaz-Pacheco A, Álvarez-Carmona MA, Guerrero-Rodríguez R, Chávez LAC, Rodríguez-González AY, Ramírez-Silva JP, Aranda R (2024) Artificial intelligence methods to support the research of destination image in tourism. A systematic review. J Exp Theor Artif Intell 36(7):1415–1445. https://doi.org/10.1080/0952813X.2022.2153276

DTI (2021) *Metodología – Red de destinos turísticos inteligentes*. https://www.destinosinteligentes.es/metodologia-dti/. Accessed 06 August 2024

Ercan F (2023) Smart tourism destination: A bibliometric review. Eur J Tourism Res 34:3409. https://doi.org/10.54055/ejtr.v34i.2788

Europarl (2025) Sparsely populated and under-populated areas. European Parliament. https://www.europarl.europa.eu/RegData/etudes/BRIE/2016/586632/EPRS_BRI(2016)586632_EN.pdf. Accessed 02 February 2025

Eurostat (2025) *Local Administrative Units (LAU)*. Eurostat. https://ec.europa.eu/eurostat/web/nuts/local-administrative-units. Accessed 02 February 2025

Faerber LS, Hofmann J, Ahrholdt D, Schnittka O (2021) When are visitors actually satisfied at visitor attractions? What we know from more than 30 years of research. Tourism Manag 84:104284. https://doi.org/10.1016/j.tourman.2021.104284

Fauzel S, Tandrayen-Ragoobur V, Matadeen SJ (2024) A Literature Survey of Technological Disruptions in the Service Sector. In: Singh Kaurav RP, Mishra V (eds) Review of Technologies and Disruptive Business Strategies. Review of Management Literature, vol 3. Emerald Publishing Limited, Leeds, pp 185–202. https://doi.org/10.1108/S2754-586520240000003010

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From Data Mining to Knowledge Discovery in Databases. AI Mag 17(3):37–54. https://doi.org/10.1609/aimag.v17i3.1230

Filieri R, Alguezaui S, McLeay F (2015) Why do travelers trust TripAdvisor? Antecedents of trust toward consumer-generated media and its influence on recommendation adoption and word of mouth. Tourism Manag 51:174–185. https://doi.org/10.1016/j.tourman.2015.05.007

Gani MO, Roy H, Faroque AR, Rahman MS, Munawara M (2024) Smart tourism technologies for the psychological well-being of tourists: a Bangladesh perspective. J Hosp Tourism Insights 7(3):1371–1390. https://doi.org/10.1108/JHTI-06-2022-0239
Garanti Z (2023) Value co-creation in smart tourism destinations. Worldw Hosp Tourism Themes 15(5):468–475. https://doi.org/10.1108/WHATT-06-2023-0070
García Castillo LA, Catalán Villegas A, Cuevas Valencia RE (2024) El Impacto de Google Maps Platforms en la Eficiencia del Sector Turístico. XIKUA Bol Cient Esc Super Tlahuelilpan 12(Especial):25–32. https://doi.org/10.29057/xikua.v12iEspecial.12716
Gatto R, Santopietro L, Scorza F (2022) Tourism and Abandoned Inland Areas Development Demand: A Critical Appraisal. In: Gervasi O, Murgante B, Misra S, Rocha AMAC, Garau C (eds) Computational Science and Its Applications – ICCSA 2022 Workshops. ICCSA 2022. Lecture Notes in Computer Science, vol 13382. Springer, Cham, pp 40–47. https://doi.org/10.1007/978-3-031-10592-0_4
Geetha M, Singha P, Sinha S (2017) Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. Tourism Manag 61:43–54. https://doi.org/10.1016/j.tourman.2016.12.022
Gelter J, Lexhagen M, Fuchs M (2021) A meta-narrative analysis of smart tourism destinations: implications for tourism destination management. Curr Issues Tourism 24(20):2860–2874. https://doi.org/10.1080/13683500.2020.1849048
George OA, Ramos CMQ (2024) Sentiment analysis applied to tourism: exploring tourist-generated content in the case of a wellness tourism destination. Int J Spa Wellness 7(2):139–116. https://doi.org/10.1080/24721735.2024.2352979
Ghazali RM, Cai L (2013) Social Media Sites in Destination Image Formation. Tourism Social Media: Transformations in Identity, Community and Culture. Tourism Social Science Series, vol 18. Emerald Group Publishing Limited, Leeds, pp 73–86. https://doi.org/10.1108/S1571-5043(2013)0000018007
Giglio S, Bertacchini F, Bilotta E, Pantano P (2019) Using social media to identify tourism attractiveness in six Italian cities. Tourism Manag 72:306–312. https://doi.org/10.1016/j.tourman.2018.12.007
Gomes S, Lopes JM, Ferreira L (2024) Looking at the tourism industry through the lenses of industry 4.0: a bibliometric review of concerns and challenges. J Hosp Tourism Insights 7(1):436–457. https://doi.org/10.1108/JHTI-10-2022-0479
Gretzel U, Werthner H, Koo C, Lamsfus C (2015) Conceptual foundations for understanding smart tourism ecosystems. Comput Hum Behav 50:558–563. https://doi.org/10.1016/j.chb.2015.03.043
Grundner L, Neuhofer B (2021) The bright and dark sides of artificial intelligence: A futures perspective on tourist destination experiences. J Destination Mark Manag 19:100511. https://doi.org/10.1016/j.jdmm.2020.100511
Hamdy A, Zhang J, Eid R (2024) A new proposed model for tourists' destination image formation: the moderate effect of tourists' experiences. Kybernetes 53(4):1545–1566. https://doi.org/10.1108/K-11-2022-1525
Hashimoto A, Telfer DJ, Telfer S (2021) Life beyond growth? Rural depopulation becoming the attraction in Nagoro, Japan's scarecrow village. J Herit Tourism 16(5):493–512. https://doi.org/10.1080/1743873X.2020.1807556
Herrera-Prado A-L, Velarde-Valdez M, Olmos-Martínez E (2024) Participación de las TIC en los modelos de Destinos Turísticos Inteligentes. Investig Tur 271–28. https://doi.org/10.14198/INTURI.20111
Hiererra SE, Meyliana M, Ramadhan A, Purnomo F (2023) The requirement aspect for sustainability smart tourism destinations: a systematic literature review and proposed model analysis. J Eng Sci Technol 18(6):2895–2914
Howie F (2003) Managing the tourist destination. Continuum, London
Hu M, Li H, Song H, Li X, Law R (2022) Tourism demand forecasting using tourist-generated online review data. Tourism Manag 90:104490. https://doi.org/10.1016/j.tourman.2022.104490
Huda C, Heryadi Y, Lukas, Budiharto W (2024) Smart Tourism Recommender System Modelling Based on Hybrid Technique and Content Boosted Collaborative Filtering. IEEE Access 12:131794–131808. https://doi.org/10.1109/ACCESS.2024.3450882
Hung YC, Yang YL, Yang HE, Chuang YH (2011) Factors Affecting the Adoption of E-commerce for the Tourism Industry in Taiwan. Asia Pac J Tourism Res 16(1):105–119. https://doi.org/10.1080/10941665.2011.539394
Hunter WC, Chung N, Gretzel U, Koo C (2015) Constructivist research in smart tourism. Asia Pac J Inf Syst 25(1):103–118. https://doi.org/10.14329/apjis.2015.25.1.103

Ivars-Baidal JA, Celdrán-Bernabeu MA, Mazón J-N, Perles-Ivars F (2017) Smart destinations and the evolution of ICTs: A new scenario for destination management? Curr Issues Tourism 22(13):1581–1600. https://doi.org/10.1080/13683500.2017.1388771

Ivars-Baidal J, Casado-Díaz AB, Navarro-Ruiz S, Fuster-Uguet M (2024) Smart tourism city governance: exploring the impact on stakeholder networks. Int J Contemp Hosp Manag 36(2):582–601. https://doi.org/10.1108/IJCHM-03-2022-0322

Januszewska M, Jaremen DE, Nawrocka E (2015) The effects of the use of ICT by tourism enterprises. Zesz Naukowe Uniw Szcz Serv Manag 16(2):65–73. https://doi.org/10.18276/smt.2015.16-07

JCyL (2025) *Boletín de Coyuntura Turística*. Portal Oficial de Turismo de la Junta de Castilla y León. https://www.turismocastillayleon.com/es/espacio-profesionales/estudios-estadisticas/boletin-coyuntura-turistica?isprediction=1. Accessed 02 February 2025

Jesus C, Franco M (2016) Cooperation networks in tourism: A study of hotels and rural tourism establishments in an inland region of Portugal. J Hosp Tourism Manag 29:165–175. https://doi.org/10.1016/j.jhtm.2016.07.005

John SP, Supramaniam S (2024) Value co-creation research in tourism and hospitality management: A systematic literature review. J Hosp Tourism Manag 58:96–114. https://doi.org/10.1016/j.jhtm.2023.11.008

Jovicic DZ (2019) From the traditional understanding of tourism destination to the smart tourism destination. Curr Issues Tourism 22(3):276–282. https://doi.org/10.1080/13683500.2017.1313203

Kalandarovna-Abdurakhmanova G, Tashtemirovich-Astanakulov O, Bakhodirovich-Goyipnazarov S, Bakhramovna-Irmatova A (2022) Tourism 4.0: Opportunities for applying industry 4.0 technologies in tourism. In: Proceedings of the 6th International Conference on Future Networks & Distributed Systems (ICFNDS'22). ACM, Tashkent, pp 33–38. https://doi.org/10.1145/3584202.358420

Kushcheva N, Eilola T-M (2023) Relevance of Social Media Management in Online Reputation Building in Tourism and Hospitality: Case of Finland. In: Rialti R, Kvítková Z, Makovník T (eds) Online Reputation Management in Destination and Hospitality. Emerald Publishing Limited, Leeds, pp 163–174. https://doi.org/10.1108/978-1-80382-375-120231009

Lasisi TT, Odei SA, Eluwole KK (2023) Smart destination competitiveness: underscoring its impact on economic growth. https://doi.org/10.1108/JTF-09-2022-0243. J Tourism Futur ahead-of-print

Lee JH, Phaal R, Lee S-H (2013) An integrated service-device-technology roadmap for smart city development. Technol Forecast Soc Chang 80(2):286–306. https://doi.org/10.1016/j.techfore.2012.09.020

Leiras A, Eusébio C (2023) Perceived image of accessible tourism destinations: a data mining analysis of Google Maps reviews. Curr Issues Tourism 27(16):2584–2602. https://doi.org/10.1080/13683500.2023.2230338

Li Y, Hu C, Huang C, Duan L (2017) The concept of smart tourism in the context of tourism information services. Tourism Manag 58:293–300. https://doi.org/10.1016/J.TOURMAN.2016.03.014

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

Li X, Law R, Xie G (2021) Review of tourism forecasting research with Internet data. Tourism Manag 83:104245. https://doi.org/10.1016/j.tourman.2020.104245

Lin Y, Huang J (2006) Internet blogs as a tourism marketing medium: A case study. J Bus Res 59(10–11):1201–1205. https://doi.org/10.1016/j.jbusres.2005.11.005

Liu J, Hall CM, Zhu C, Ting Pong Cheng V (2023) Redefining the concept of smart tourism in tourism and hospitality. Anatolia 35(3):566–578. https://doi.org/10.1080/13032917.2023.2282712

Lyu J, Khan A, Bibi S, Chan JH, Qi X (2022) Big data in action: An overview of big data studies in tourism and hospitality literature. J Hosp Tourism Manag 51:346–360. https://doi.org/10.1016/j.jhtm.2022.03.014

Mandić A, Garbin-Praničević D (2019) Progress on the role of ICTs in establishing destination appeal: Implications for smart tourism destination development. J Hosp Tourism Technol 10(4):791–813. https://doi.org/10.1108/JHTT-06-2018-0047

Mariani M, Baggio R (2022) Big data and analytics in hospitality and tourism: a systematic literature review. Int J Contemp Hosp Manag 34(1):231–278. https://doi.org/10.1108/IJCHM-03-2021-0301

Marine-Roig E, Huertas A (2020) How safety affects destination image projected through online travel reviews. J Travel Res 59(4):640–659. https://doi.org/10.1177/0047287519855130

Martín-Galán CA (2020) *Diseño e implementación de un Sistema Automático para la detección de opiniones turísticas en Redes Sociales*. Doctoral dissertation, Universidad de La Laguna. https://www.educacion.gob.es/teseo/imprimirFicheroTesis.do?idFichero=mEdgnvkcXNQ%3D. Accessed 11 December 2024

Martínez-Plumed F, Ramírez-Quintana MJ, Kull M, Ferri C, Lachiche N, Hernandez-Orallo J, Contreras-Ochando L, Flach P (2020) Data Science Trajectories. IEEE Trans Knowl Data Eng 33(8):3048–3061. https://doi.org/10.1109/tkde.2019.2962680. CRISP-DM Twenty Years Later: From Data Mining Processes to

Meijomil S (2022) *Reputación online: qué es, su importancia y cómo cuidarla*. https://www.inboundcycle.com/blog-de-inbound-marketing/reputacion-online-guia-completa. Accessed 10 September 2024

Mohammadi F, Yazdani HR, Jami Pour M, Soltani M (2021) Co-creation in tourism: a systematic mapping study. Tourism Rev 76(2):305–343. https://doi.org/10.1108/TR-10-2019-0425

Moro S, Rita P, Vala B (2016) Predicting social media performance metrics and evaluation of online reputation: A data mining approach. J Bus Res 69(9):3341–3351. https://doi.org/10.1016/j.jbusres.2016.02.010

Mukherjee S, Zalani C (2024) 29 Google Maps StatisTIC: Verified and updated for 2025 - On the Map Internet Marketing. https://www.onthemap.com/blog/google-maps-statisTIC/. Accessed 10 January 2025

Nolasco-Cirugeda A, García-Mayor C, Lupu C, Bernabeu-Bautista A (2022) Scoping out urban areas of tourist interest though geolocated social media data: Bucharest as a case study. Inf Technol Tourism 24(3):361–387. https://doi.org/10.1007/s40558-022-00235-8

O'Connor P (2023) Small- and medium-sized tourism enterprises and smart tourism: Tourism Agenda 2030 perspective article. Tourism Rev 78(2):339–343. https://doi.org/10.1108/TR-09-2022-0431

Pan B, MacLaurin T, Crotts JC (2007) Travel Blogs and the Implications for Destination Marketing. J Travel Res 46(1):35–45. https://doi.org/10.1177/0047287507302378

Pan X, Rasouli S, Timmermans H (2021) Investigating tourist destination choice: Effect of destination image from social network members. Tourism Manag 83:104217. https://doi.org/10.1016/j.tourman.2020.104217

Paolanti M, Mancini A, Frontoni E, Felicetti A, Marinelli L, Marcheggiani E, Pierdicca R (2021) Tourism destination management using sentiment analysis and geo-location information: a deep learning approach. Inf Technol Tourism 23:241–264. https://doi.org/10.1007/s40558-021-00196-4

Pártlová P, Dušek R, Sagapova N (2022) Building reputation and social media–how effectively do attractive European tourist destinations communicate on them? Entrepreneurship Sustain Issues 10(1):467. https://doi.org/10.9770/jesi.2022.10.1(26)

Peceny US, Urbančič J, Mokorel S, Kuralt V, Ilijaš T (2019) Tourism 4.0: Challenges in marketing a paradigm shift. Consumer behavior and marketing. IntechOpen, London, pp 1–19. https://doi.org/10.5772/intechopen.84762

Pencarelli T (2020) The digital revolution in the travel and tourism industry. Inf Technol Tourism 22(3):455–476. https://doi.org/10.1007/s40558-019-00160-3

Platov AV, Litvinova OI, Troitskaya NG (2020) *The structure of the online reputation analysis in Tourism*. In: Proceedings of the 2nd International Scientific and Practical Conference Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth (MTDE 2020). Atlantis Press, Yekaterinburg, pp 813–816. https://doi.org/10.2991/aebmr.k.200502.132

Plotnikova V, Dumas M, Milani F (2020) Adaptations of data mining methodologies: a systematic literature review. PeerJ Comput Sci 6:e267. https://doi.org/10.7717/peerj-cs.267

Pollák F, Dorčák P, Račeta N, Svetozarovová N (2016) Sustainable e-marketing of selected tourism subjects from the Mediterranean through active online reputation management. In: Leon-Garcia A et al (eds) Smart City 360°. SmartCity 360 SmartCity 360 2016 2015. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 166. Springer, Cham, pp 692–703. https://doi.org/10.1007/978-3-319-33681-7_60

Prat Forga JM (2020) La puesta en valor turístico del patrimonio cultural en fase de desaparición. Una oportunidad para los territorios de interior. Los casos de los monasterios de Vilabertrán, Scala Dei y Escornalbou, en Catalunya. Investig Tur 20:315–334. https://doi.org/10.14198/INTURI2020.20.14

Qin Y, Wang X, Xu Z (2022) Ranking tourist attractions through online reviews: A novel method with intuitionistic and hesitant fuzzy information based on sentiment analysis. Int J Fuzzy Syst 24(2):755–777. https://doi.org/10.1007/s40815-021-01131-9

Reichheld FF (2003) The one number you need to grow. Harv Bus Rev 81(12):46–55

Ribes JFP, Baidal JI (2018) Smart sustainability: A new perspective in the sustainable tourism debate. Investig Region-J Region Res 42:151–170

Rong J, Hao H, Xu W (2024) Big data intelligent tourism management platform design based on abnormal behavior identification. Intell Sys Appl 21:200312. https://doi.org/10.1016/j.iswa.2023.200312

Samala N, Katkam BS, Bellamkonda RS, Rodríguez RV (2022) Impact of AI and robotics in the tourism sector: a critical insight. J Tourism Futur 8(1):73–87. https://doi.org/10.1108/JTF-07-2019-0065
Saranow J (2004) Getting Travel Advice From a Stranger; Online Blogs Emerge As Popular Resource; How to Find Useful Ones. Wall Str J 28(3)
Scorza F, Pilogallo A, Las Casas G (2019) Investigating Tourism Attractiveness in Inland Areas: Ecosystem Services, Open Data and Smart Specializations. In: Calabrò F, Della Spina L, Bevilacqua C (eds) New Metropolitan Perspectives. ISHT 2018. Smart Innovation, Systems and Technologies, vol 100. Springer, Cham, pp 30–38. https://doi.org/10.1007/978-3-319-92099-3_4
SEGITTUR (2020) *Destinos turísticos inteligentes - SEGITTUR*. https://www.segittur.es/destinos-turisticos-inteligentes/proyectos-destinos/destinos-turisticos-inteligentes/. Accessed 06 August 2024
Serra P, Pallares-Barbera M, Salvati L (2022) Can a long-term economic subsidy counteract rural depopulation? An empirical analysis from two spanish regions. Qual Quant 56(4):1–20. https://doi.org/10.1007/s11135-022-01505-7
Shafiee S, Ghatari AR, Hasanzadeh A, Jahanyan S (2019) Developing a model for sustainable smart tourism destinations: A systematic review. Tourism Manag Perspect 31:287–300. https://doi.org/10.1016/j.tmp.2019.06.002
Shafiee S, Rajabzadeh Ghatari A, Hasanzadeh A, Jahanyan S (2021) Smart tourism destinations: a systematic review. Tourism Rev 76(3):505–528. https://doi.org/10.1108/TR-06-2019-0235
Shafiee S, Ghatari AR, Hasanzadeh A, Jahanyan S (2022) Developing a model for smart tourism destinations: an interpretive structural modelling approach. Inf Technol Tourism 24(4):511–546. https://doi.org/10.1007/s40558-022-00236-7
Sigala M (2018) New technologies in tourism: From multi-disciplinary to anti-disciplinary advances and trajectories. Tourism Manag Perspect 25:151–155. https://doi.org/10.1016/j.tmp.2017.12.003
Stankov U, Gretzel U (2020) Tourism 4.0 technologies and tourist experiences: a human-centered design perspective. Inf Technol Tourism 22(3):477–488. https://doi.org/10.1007/s40558-020-00186-y
STD (2024) Smart Tourism Destinations. European Union. https://smarttourismdestinations.eu/about-us/. Accessed 21 December 2024
Sun S, Wei Y, Tsui KL, Wang S (2019) Forecasting tourist arrivals with machine learning and internet search index. Tourism Manag 70:1–10. https://doi.org/10.1016/j.tourman.2018.07.010
Sunagar P, Hanumantharaju R, Siddesh GM, Kanavalli A, Srinivasa KG (2020) Influence of big data in smart tourism. In: Bhattacharyya S, Snášel V, Gupta D, Khanna A (eds) Hybrid Computational Intelligence. Academic, Cambridge, pp 25–47. https://doi.org/10.1016/B978-0-12-818699-2.00002-0
Sustacha I, Baños-Pino JF, Del Valle E (2023) The role of technology in enhancing the tourism experience in smart destinations: A meta-analysis. J Destination Mark Manag 30:100817. https://doi.org/10.1016/j.jdmm.2023.100817
Tavitiyaman P, Qu H, Tsang WSL, Lam CWR (2021) The influence of smart tourism applications on perceived destination image and behavioral intention: The moderating role of information search behavior. J Hosp Tourism Manag 46:476–487. https://doi.org/10.1016/j.jhtm.2021.02.003
Tham A, Mair J, Croy G (2019) Social media influence on tourists' destination choice: importance of context. Tourism Recreat Res 45(2):161–175. https://doi.org/10.1080/02508281.2019.1700655
Tripathi S, Emmert-Streib F, Jodlbauer H, Dehmer M, Muhr D, Brunner M (2021) Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing. Front Artif Intell 4:576892. https://doi.org/10.3389/frai.2021.576892
UNE (2022) UNE 178502:2022. *Indicadores y herramientas de los destinos turísticos inteligentes*. https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0068105. Accessed 06 August 2024
Urrutia XAM, Urrutia DCM, Larenas PC, Urrutia EKM, Ocaña JM (2018) The french way of St. James as an engine in the invigoration of rural municipalities in Galicia. Rev Econ Sociol Rural 56(3):425–436. https://doi.org/10.1590/1234-56781806-94790560304
Vecchio P, Del, Mele G, Ndou V, Secundo G (2018) Creating value from Social Big Data: Implications for Smart Tourism Destinations. Inf Process Manag 54(5):847–860. https://doi.org/10.1016/j.ipm.2017.10.006
Vianello D, Marrucci A, Ciappei C, Becagli C (2023) Big Data and Online Reputation Management in Tourism: Leveraging the Role of Entrepreneurship. In: Rialti R, Kvítková Z, Makovnik T (eds) Online Reputation Management in Destination and Hospitality: What We Know, What We Need to Know. Emerald Publishing Limited, Leeds, pp 73–90. https://doi.org/10.1108/978-1-80382-375-120231004

Vidal-Matzanke A, Vidal-González P (2022) Hiking accommodation provision in the mountain areas of Valencia Region, Spain: a tool for combating the depopulation of rural areas. J Sport Tourism 26(2):165–184. https://doi.org/10.1080/14775085.2022.2053563
Wang J, Li Y, Wu B, Wang Y (2021) Tourism destination image based on tourism user generated content on internet. Tourism Rev 76(1):125–137. https://doi.org/10.1108/TR-04-2019-0132
Wei W, Önder I, Uysal M (2024) Smart tourism destination (STD): developing and validating an impact scale using residents' overall life satisfaction. Curr Issues Tourism 27(17):2849–2872. https://doi.org/10.1080/13683500.2023.2296587
Wenger A (2008) Analysis of travel bloggers' characteristics and their communication about Austria as a tourism destination. J Vacat Mark 14(2):169–176. https://doi.org/10.1177/1356766707087525
Wu DC, Zhong S, Song H, Wu J (2024) Do topic and sentiment matter? Predictive power of online reviews for hotel demand forecasting. Int J Hosp Manag 120:103750. https://doi.org/10.1016/j.ijhm.2024.103750
Wu DC, Zhong S, Wu J, Song H (2025) Tourism and Hospitality Forecasting With Big Data: A Systematic Review of the Literature. J Hosp Tourism Res 49(3):615–634. https://doi.org/10.1177/10963480231223151
Xiang Z, Tussyadiah I, Buhalis D (2015) Smart destinations: Foundations, analytics, and applications. J Destination Mark Manag 4(3):143–144. https://doi.org/10.1016/j.jdmm.2015.07.001
Xiang Z, Du Q, Ma Y, Fan W (2017) A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. Tourism Manag 58:51–65. https://doi.org/10.1016/j.tourman.2016.10.001
Xu H, Lv Y (2022) Mining and application of tourism online review text based on natural language processing and text classification technology. Wirel Commun Mob Comput 2022(1):9905114. https://doi.org/10.1155/2022/9905114
Zeng B, Gerritsen R (2014) What do we know about social media in tourism? A review. Tourism Manag Perspec 10:27–36. https://doi.org/10.1016/j.tmp.2014.01.00

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

**Julio César Puche-Regaliza[1] · Isabel Marcilla-Lombraña[1] · Paula Antón-Maraña[1] · Pablo Arranz-Val[1]**

✉ Paula Antón-Maraña
panton@ubu.es

Julio César Puche-Regaliza
jcpuche@ubu.es

Isabel Marcilla-Lombraña
iml1009@alu.ubu.es

Pablo Arranz-Val
parranz@ubu.es

[1] University of Burgos, Burgos, Spain