



How to exploit Big Social Data in the Covid-19 pandemic: the case of the Italian tourism industry

Gianluca Solazzo¹ · Ylenia Maruccia¹  · Valentina Ndou¹ · Pasquale Del Vecchio¹

Received: 20 October 2020 / Accepted: 25 February 2022 / Published online: 12 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

From March 2020, restrictive measures have been taken to limit the Covid-19 pandemic, thus impacting and hugely affecting the tourism industry. In this new scenario, tourism destination managers and stakeholders need to look ahead to the future, trying to wisely react to overcome this crisis. This paper aims at demonstrating how a Smart Tourism Destination can tackle the consequences of the Covid-19 pandemic using smart and digital technologies such as Big Data and Analytics. In particular, the monitoring of Italian public debates on social media allowed discovering knowledge from unstructured data that provide valuable insights for destination stakeholders and decision-makers.

Keywords Covid-19 · Big Social Data · Smart Tourism Destination · Topic extraction · Sentiment analysis

1 Introduction

In the twenty-first century, international tourism has been affected by a series of disruptive events, such as health-related crises (e.g. the severe acute respiratory syndrome (SARS) outbreak in 2003, avian influenza “bird flu”—Influenza A(H1N1)—in 2009 and the Middle East Respiratory Syndrome (MERS) outbreak in 2015), the

✉ Ylenia Maruccia
ylenia.maruccia@unisalento.it

Gianluca Solazzo
gianluca.solazzo@unisalento.it

Valentina Ndou
valentina.ndou@unisalento.it

Pasquale Del Vecchio
pasquale.delvecchio@unisalento.it

¹ Department of Engineering for Innovation, University of Salento – ‘Building “Aldo Romano”, Campus Ecotekne, Lecce, Italy

worldwide economic crisis in 2008/2009 (Gössling et al. 2020), several natural disasters (e.g. Hurricane Katrina in the US in 2005 and the Indian Ocean tsunami in 2004) and terrorist attacks (e.g. Charlie Hebdo and Bataclan in Paris in 2015, the Christmas markets in Berlin in 2016, the Champs-Élysées in Paris in 2017, the Ramblas in Barcelona and Cambrils in 2017). However, none of these has had an impact such as the one registered with the Covid-19 pandemic outbreak in 2020.

The coronavirus (Covid-19) pandemic is an unprecedented event that is causing huge damage to the global tourism industry and incomparable with all the aforementioned disasters or crises. The Covid-19 pandemic has been an event that broke out suddenly, without any kind of warning, and can be clearly defined as “a sudden unpredictable catastrophic change over which it has little control” (Faulkner 2001). Therefore, the usual prevention and planning phases of disaster prevention could not be implemented. According to the Worldometer statistics, on October 13th, 2020 the number of Covid-19-related deaths has surpassed 1,000,000, and this has created widespread fear of death in people. Due to the coronavirus worldwide spreading, most countries have taken restrictive measures with the final aim of trying to limit the pandemic and control the number of infected people. Borders have been closed, all air flows were suspended, holidays and cruises were cancelled, and hotels, restaurants, and tourist attractions were closed (Zenker and Kock 2020). These whole significant precautionary measures to contain Covid-19 diffusion have had clear impacts on the economy on an international scale (Iacus et al. 2020). Moreover, the global diffusion of Covid-19 has had not only an impact on the tourism sector but has provoked changes in consumers’ needs and market demands that were satisfied by traditional business models (Gössling et al. 2020).

The studies of the impacts of crisis events and calamities in the tourism sector and how this industry and the relevant government agencies seek to find responses are particularly important. Before the Covid-19 pandemic, researchers have achieved moderate progress concerning the themes of designing and implementing crisis recovery and response strategies in the tourism industry (McKercher and Chon 2004), studying also how to be able to recover quickly from difficulties to address future crises (Hall et al. 2017). Although Faulkner (2001) contends that tourism is as inclined to calamities as any other sector because it is an area of human activity, Cavlek (2002) demonstrated that recovering tourism from a crisis have resulted in more troublesome than other sectors. The Covid-19 pandemic made tourism stakeholders face a completely new and unpredictable situation, in which any certainty and knowledge acquired in the past were no longer safe. At the same time, the structural break caused by the Covid-19 pandemic raised the need for appropriate policy, together with a huge allocation of resources, and governments, national tourism organizations, foreign tour operators, local travel organizers and local hospitality officials needed to establish a very strong partnership between them (Cavlek 2002) to ensure the realization of numerous significant actions. These actions consist of an effective reconstruction of the destination image, the necessity to overcome any negative policy consequential from the crisis, the way to restore in the short-term and reconstruct in the long-term all the tourism facilities and infrastructures which have suffered the consequences of the crisis, how to effectively manage media coverage,

how to reduce barriers and facilitate travel, and business and consumer regulation support and subsidies (Steiner et al. 2012; Cavlek 2002).

Nevertheless, the knowledge of how the crisis can encourage changes in the sector, how companies can ride the wave of the crisis and reconvert difficulties into transformative innovation, how to derive innovative insights for transformation management and sustainable perspective (UNWTO 2020a), how to enable, inform and reshape thought and restoration of upcoming normality (Sigala 2020) is still missing.

In this perspective, becoming a Smart Tourism Destination (hereinafter STD), adopting and exploiting smart technologies can help to minimize the aforementioned knowledge gaps while limiting the consequences of the spread of Covid-19 (Sigala 2020). A STD refers to the intensive use of technology for supporting the destinations to enrich the tourist experience, implement innovative technologies that facilitate access to services and provide faster information sharing, respond to the needs of stakeholders (tourists, tourism institutions and organizations, government, citizens, etc.), minimize the use of resources, pursue a sustainable society and destination, citizens to improve the quality of life and the quality of the travel experience for visitors (Duran and Uygur 2022). This concept foresees ICT-based tools and other digital technologies (such as Big Data and Internet of Things) being used for data collection, processing and analysing, to ensure more efficient and effective use of resources, higher satisfaction of tourists, as well as the well-being of residents. In particular, Big Data and Analytics techniques face unprecedented challenges and opportunities post COVID-19 to build on the huge amount of data and a new tourism model that is more sustainable, smarter, and safer than those previously implemented (Montero and López-Sánchez 2021).

By March 9th 2020, Italy had officially shown cases of Covid-19. To handle these serious circumstances and to decrease the diffusion of the infections, a series of decrees imposed restrictions on the movement of individuals in the whole national territory. People have been allowed to leave their homes only for limited and documented purposes and many activities have been temporarily locked. During the first lockdown, social media platforms have offered the opportunity to enhance social isolation, allowing people to communicate without time and space limits, and discussions have hugely moved from the real world to the virtual one. Social media has emerged as an important medium for governments to explain the pandemic situation and communicate political decisions, and for tourism stakeholders and citizens to speak about crisis, comment on public decisions, and express their views and concerns about the possible consequences of the lockdown accordingly. Thus, this huge usage of social media has led to an enormous growth of users' generated data, which provide an awesome opportunity to mine valuable insights and better understand worries, anxieties and business interests to slow down the impacts of the pandemic crisis. This has enabled on the one hand the development of smart governance, ensuring a collaborative, participatory and communication-based environment, and, on the other hand, the continuous collaboration between enterprises, local and national administrations, tourists and residents, which are some theoretical pillars for the STDs development.

Italy officially entered the Second Phase on May 4th 2020, characterized by several actions aiming at alleviating the Covid-19 impact and finding some ways to recover the drastic losses of the industry. These measures and actions, undertaken by the government, were at the centre of public debates on social media (Pasquinelli and Trunfio 2021).

Based on the above premises, this paper aims at contributing to the research on the integration of resilience thinking into operational and decision-making settings of a STD with the final purpose to tackle the consequences of the Covid-19 pandemic. The methodological framework developed in this study, by exploiting Big Social Data and Analytics through the monitoring of online discussions, aims at giving insights to businesses, governments and tourism operators for addressing the restart of tourism. In particular, this paper tries to answer the following research question:

How a tourism destination can tackle the consequences of the Covid-19 pandemic using Big Social Data and Analytics insight from public debate over social media?

Public debate on Twitter has been monitored in the period from 13th April to June 15th 2020 to understand how the Italian tourism industry have tried to react to this crisis during the Second Phase, and eventually highlighting tourism stakeholders' drivers, actions and reactions to the Covid-19 impacts (Sigala 2020). This approach can lead both tourism stakeholders and the tourism destinations, whose activities are affected by this pandemic, to shift towards a STD paradigm, where capturing and analysing Big Social Data (hereinafter BSD) generated by tourists is a useful instrument that gives useful insights, nourishes the value creation process within a STD (Del Vecchio et al. 2018), helps to manage the crisis and ameliorates the tourists experience.

This paper is structured as follows. In Sect. 2, the related literature background is presented, while in Sect. 3 the methodology adopted for the analysis has been described. Section 4 presents the results obtained from the analysis, while in Sect. 5 discussions on the achieved results are presented. Finally, the originality of this paper, limitations, implications and future research are discussed in Sect. 6.

2 Literature review

2.1 Smart tourism destinations

Information and Communication Technologies (ICTs) currently have a prominent role in any area of society, in particular, they are important in all the activities making up the tourist sector (Gretzel et al. 2016). The digital-technological revolution has deeply altered tourism management (Gretzel et al. 2000) and the impacts of ICTs in the tourism sector emphasize the importance of investing in companies that provide tourist services. According to Sigalat-Signes et al. (2020), these investments in ICTs should allow tourist destinations to refine their competitive advantage and improve their management. Focusing on this context, characterized by the use

of ICTs and technological evolution, the adoption of new ideas and approaches for tourism development enables new services and the conversion of traditional ones in the sector.

These technological advancements in tourism have led to the birth of the STD notion, where ICT is a driver for smart growth and the competitiveness of a destination (Femenia-Serra and Ivars-Baidal 2018). The STD is understood as a local tourism system characterized by high innovation and the presence of both advanced services and open, integrated and shared processes, useful for improving the quality of life for both local people and tourists (Micera et al. 2013; Wang et al. 2013; Caragliu et al. 2013).

Smart Tourism counts on extensive adoption of advanced digital technologies and applications, such as social media sensors, the Internet of things, smart devices and sensors to collect and exploit the huge amount of data for creating new value propositions (Gretzel et al 2015; Sigala et al. 2012). Thanks to all these technologies tourists are highly engaged in producing content (User Generated Content) and leaving digital traces during all phases of their trip, when they plan, when they consume or when they turn back, by sharing their emotions and providing feedback on their experience in social media platforms or through online surveys (Hu et al. 2017; Wang et al. 2013; Fuchs and Höpken 2011). Therefore, the foremost driver for smart destination competitiveness relies heavily on the abilities of destination managers to collect and aggregate this large amount of data and intelligently exploit them for creating value into competitive assets (Del Vecchio et al. 2018).

The process to become a STD is very intricate and challenging (Femenia-Serra et al. 2019) and involves different perspectives in terms of models, tools and strategies (Del Vecchio et al. 2018). In particular, to become a STD means to have a participatory government that will ensure political strategies and policies aimed at enhancing sustainable development and economic growth in the tourism sector, to have available advanced ICT infrastructure, to have access to real-time information (coming from sensors or residents and visitors as digital footprint from their social media activities), to digitalize the core of business processes, the tourism structure to adopt technologies and improve their services, to engage both the community and the tourist participation for co-creating tourism experience, and so on (Maruccia et al. 2019).

A STD goal is to develop and deliver a smart, personalized, context-aware and real-time experience, by interconnecting different stakeholders, through dynamic platforms, knowledge-intensive communication flows and enhanced decision support systems (Buhalis and Amaranggana 2015).

Social media and advanced technologies are influential in empowering STDs to develop such dynamic connections and to benefit from a vast source of tourist information knowledge and opinions realized through conversational media (Miah et al. 2017; Buhalis and Foerste 2015).

2.2 Smart tourism destination and Big Data

The concept of the smart destination is part of the evolutionary concept of the smart city (Trunfio and Pasquinelli 2021), and it intertwines smart technological tools, people and institutions to create public value (Desdemoustier et al. 2019). A STD can be seen as the result of the interconnection of tourism destinations with multiple stakeholders' communities through dynamic platforms and knowledge-intensive flows of communication and enhanced decision support systems (Jovicic 2019; Buhalis and Amaranggana 2015). While a fundamental purpose of a STD consists of stimulating and facilitating the highest tourist satisfaction and experience (García-Milon et al. 2020), optimizing both the destination's competitiveness and consumer satisfaction (Del Vecchio et al. 2018), the role of ICTs, within the STD, is to provide the platforms through which knowledge and information are instantly and easily exchanged, facilitating stakeholders' collaboration (Jovicic 2019).

Between all the smart technologies supporting a STD, Big Data and Analytics emerged as the paradigm that is re-shaping the theory and practice of tourism (Ardito et al. 2019). The term Big Data refers to the generation and exploitation of a massive and varied amount of data from which it is possible to obtain precious insights (Abdar et al. 2017), characterized by high-volume, high-velocity and high-variety (Davenport 2013). A subset of Big Data consists of BSD, which has been defined in Solazzo et al. (2021a) as data *“generated from people’s actions and interactions within social media services and platforms, sharing a subset of Big Data properties, that needs to be collected and analysed through specific technology and analytics to provide crucial insights into human behaviour, people’s preferences and relationships, social interactions and transformations, and real-life event outcomes.”*

Social media constitute the major source of Big Data. The extant literature has widely argued the advantages of mining social media and consumer-generated content for value as well as for gaining relevant insights and information on customers' experiences, feelings, interests, opinions, behaviours, preferences, etc. (Raguseo et al. 2017; Marine-Roig and Clave 2015).

One of the major sources of BSD is social media. Social media platforms collect a huge amount of data regarding their users, both directly, through their profiles, or indirectly, through all the information generated from their actions and interactions within the platform, such as image upload, check-in in a location, or writing posts or reviews (Del Chiappa and Baggio 2015). A BSD taxonomy is presented in Olshannikova et al. (2017), in which they present the four core categories identified for BSD and synthesized as it follows: Digital Self-Representation data, i.e. data related to identifying depiction and communicative body in a digital environment; Technology-Mediated Communication data, i.e. data related to two-way communication, knowledge creation and distribution through technology; Digital Relationship Data, i.e. data that reveal digital social relationship patterns; and Digital Context Data (Solazzo et al. 2021a), i.e. data that reveal the dynamic patterns in the digital environment that constitute the users' context. In particular, the last category contains all data representing *“individual’s interests, preferences, mood and opinions, expressed in textual or multimedia format and used to communicate specific*

individual or group behaviours, social interactions, and tendencies for a wide range of social-related insights.”

In data-intensive domains, such as tourism, it is possible to properly consider them as useful resources for user modelling (Del Chiappa and Baggio 2015) and value creation processes (Höpken et al. 2019; Del Vecchio et al. 2018). Many research on Big Data and Analytics in tourism had promising applications in building a tourism recommender (Menk et al. 2017); promoting tourism (Park et al. 2016); developing smart services for urban tourism (Brandt et al. 2017); working on the identification of cultural heritage resources from geo-tagged social media (Nguyen et al. 2017); gaining relevant insights and information on customers' experiences, feelings, interests, opinions, behaviours and preferences (Kim et al. 2019; Raguseo et al. 2017; Marine-Roig and Clave 2015; Xiang et al. 2015a; Xiang et al. 2015b); supporting destination managers to enhance the destination attractiveness, shape new marketing and communication strategies, and plan tourist demand (Solazzo et al. 2021b).

Moreover, Big data have the potential to improve tourism policies and management (Xie et al. 2021; Iorio et al. 2020; Chun et al. 2020) and this evidence becomes even more crucial when considering that tourist destinations need effective management strategies and recovery measures, during and after crises, to realize sustainable development.

2.3 STD paradigm in crisis management contexts

Among the most critical events that influenced tourism destination development, the 2001 terrorist attacks, the severe acute respiratory syndrome (SARS) epidemic in 2003, the 2008–2009 world economic crisis and the Covid-19 pandemic in 2020, are classified by Taleb (2007) as black swans (unpredictable), all of which brought the world to a temporary standstill. In particular, Covid-19 has been causing unprecedented disruption for society, the economy and governments worldwide, while social distancing has forced the world to embrace digital technologies (Sułkowski 2020) as a means to facilitate a quick response to lockdown and government restrictions (Kirk and Rifkin 2020).

In such a crisis context, digital and smart technologies have contributed to a slow economic recovery of the tourism industry (Sigala 2020; World Economic Forum 2020; Buhalis 2019). In fact, they have been embraced by STD (but also by hotels, museums, parks and other tourist attractions) for creating online tourism experiences and traveller support groups and, to capture potential future customers, for using online promotional techniques in a “try before you buy” type of scenario (Aldao et al. 2021; Forbes 2020).

Indeed, the literature has already explored the use of STD implementations to enable crisis management (Gretzel and Scarpino-Johns 2018). STDs are very well suited for crisis and post-crisis recovery contexts (Bulchand-Gidumal 2022), as STDs allow for knowing where tourists are and communicating with them (Ordoñez de Pablos et al. 2015; Schroeder et al. 2013), identifying available resources than can be allocated if necessary by mechanisms such as sharing accommodation options

(Hajibaba et al. 2017), and enabling adaptive governance by choosing the right combination of top-down and participatory governance as required at each step (Gretzel et al. 2018; Lalicic and Önder 2018). Moreover, STDs can be central in helping the government to gain first-hand information about epidemics (Jia et al. 2012) or in sharing real-time information in different response stages and support local governments' decision-making (Aldunce et al. 2015). Furthermore, Big Data and Analytics have helped tourists stuck in travelling countries to come back to their home countries with the travel restriction in the Covid-19 pandemics (Kushwaha et al. 2021).

Finally, the link between STDs and crisis management has been reinforced thanks to the "Destination Resilience" concept development (Prayag 2018), based on an emerging area of research, that aims at supporting tourism managers and policy-makers in the development of more adaptive strategies in the face of vulnerabilities, growing risks and the uncertainty of crises and disasters (Bethune et al. 2022). Gretzel and Scarpino-Johns (2018) modelled a five-pillar framework of smart destination resilience, suggesting "smart tourism infrastructure and governance to equip smart destinations with sensing, opening, sharing, governing, and innovating capacities that can enhance destination resilience by supporting six specific resilience conditions". From this perspective, even if the potential of BSD for STDs management is enormous (Li et al. 2018) and its real application is still limited, it has been considered essential in crisis-related communication (Yu et al. 2021) as user-generated content is among the most trusted form of information for travellers (Filiari 2016). Recent research has explored the role of online debates from social media in providing insightful narratives that influence decision-makers and enable them to frame tourism issues during the pandemic (Pasquinelli and Trunfio 2021, 2020). In this perspective, Big Social Data and Analytics can help STDs: in decision-making on how to manage the crisis during real-time situations (Sutton et al. 2008; Vieweg et al. 2008), as they help to collect highly localised information to generate valuable insights within the limited context of a STD; in decision-making and coordination during crises (MacEachren and Cai 2006), enhancing the presentations of multidimensional highly complex data, through the sequential storage of large amounts of information in easily recognisable and searchable databases (Sigala and Marinidis 2012; Malizia et al. 2011).

This study is designed to develop a methodological framework based on the exploitation of Big Social Data and Analytics that support STD stakeholders to integrate resilience thinking into operational and decision-making settings with the final aim to tackle the consequences of the Covid-19 pandemic.

3 Methodology

To perform the analysis on the public debate generated on social media associated with the impacts of Covid-19 on the tourism sector, the following framework has been adopted (Fig. 1). The methodological framework is conceived to gain value from data generated on social media, through five layers, each of which specifies how unstructured data can be easily converted into profitable value. Thus, the methodological framework provides instructions related to the data that could be

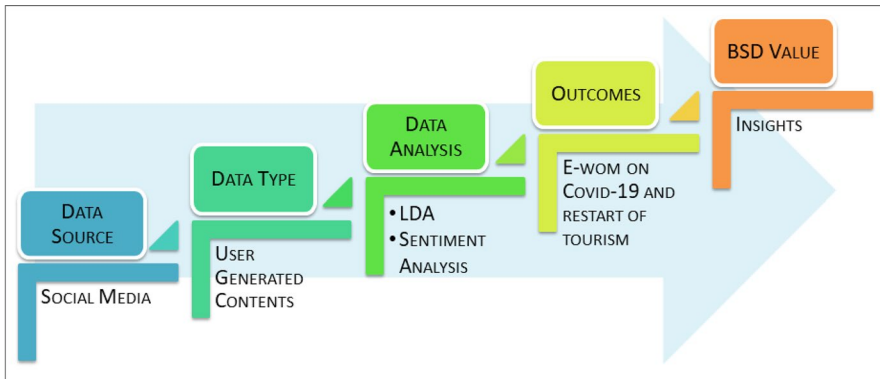


Fig. 1 The methodological framework adopted in this work

gathered to be analysed, the analysis to be performed and how their combination allows obtaining useful outcomes for extracting fruitful value.

The final aim is to find out implications and possible ideas or embryonic solutions arising from discussion to be adopted for restarting tourism, thus providing useful insights to governments, destination managers and tourism stakeholders for developing adequate policy responses.

The adoption of this framework is functional to give rigour to our analysis and determine a method for researchers to create data useful for analysis (Hesse-Biber and Leavy 2011; Carter and Little 2007). In fact, according to Avis (2003), a ‘methodological justification’ should be provided by researchers through the adoption of a methodological framework, discussing the reasons why they chose a specific method for their studies (Liamputtong and Ezzy 2005).

The first layer of the proposed framework is *Data Source*. In this case, it is composed of social media, which provide opportunities to conveniently collect large volumes of data that can be analysed. For this analysis, it has been decided to use Twitter, by adopting the combination of the following words in the search query:

search words = (#turismo AND #COVID-19) OR (#turismo AND #FaseTre) OR (#turismo AND #COVID19italia) OR (#turismo AND #coronavirus) OR (#turismo AND #coronavirusitalia) OR (#turismo AND #iorestoacasa) OR (#turismo AND #andratuttobene) OR (#turismo AND #Fase3).

and collecting 3559 tweets, from April 13th to June 15th, 2020.

The second layer is *Data Type* and it is generally made of the User Generated Contents, i.e. any form of content, such as text, images, videos and audio, that have been shared by users on online platforms. These contents should be treated when they are downloaded, to obtain “clean” data to be directly processed in the next layer. In fact, for this analysis, special characters and links have been removed from texts, to have a clean dataset.

The third layer, *Data Analysis*, is composed of different techniques that have been used to analyse the UGC. In particular:

- Latent Dirichlet Allocation (hereafter LDA), which gives support in the identification of the key topics, their meaning, how much they are prevalent in the discussion and how they relate to each other;
- Sentiment Analysis (hereafter SA), which is the interpretation and classification of emotions within text data, usually in positive, neutral and negative classes.

The fourth layer is made up of *Outcomes*, i.e. all the considerations and insights that come from the analysis. Thus, this layer serves as the basis for the last layer, *BSD Value*, providing useful insights on the online debate during the Covid-19 Second Phase and helps to detect both the sentiment and semantic meanings within the debate, understanding what are the key topics within the debate and how these opinions are polarized.

3.1 Latent Dirichlet allocation

LDA is a widespread approach for the discovery of latent topics that use multinomial probability distributions over terms, generated by soft clustering of words based on document co-occurrence. LDA gives support in the identification of the key topics, their meaning, how much they are prevalent and how they relate to each other by using three metrics: *distinctiveness*, *saliency* (Chuang et al. 2012) and *relevance* (Sievert and Shirley 2014). For the first two metrics, *distinctiveness* and *saliency*, the Kullback-Liebler divergence between the distribution of topics is computed, to measure how much a term is shared across topics—*distinctiveness*; weighting *distinctiveness* by the term's overall frequency, *saliency* is obtained (Chuang et al. 2012). Since these two metrics alone are global properties of terms, the third metric—the *relevance*, has been introduced by Sievert and Shirley (2014) to allow a deep investigation and analysis of every single topic through the visualization of several sets of terms (Sievert and Shirley 2014). In particular, this measure depends on the parameter λ , which allow obtaining a list of the most relevant terms, ranked in decreasing order and based on their probability within a specific topic ($\lambda = 1$) or within the whole dataset ($\lambda = 0$).

In general, there is not a predefined value of λ , and intermediate values of it usually are preferred as they allow to weigh the two frequency components, to extract for each cluster the most relevant terms and, finally, to evaluate the final topic using thematic considerations. Moreover, when extracting a topic from a document, there is not a predefined number of clusters to be extracted but several experiments should be done to understand what is the optimal choice for the representation of the whole dataset. In general, a good rule of thumb is to evaluate if each topic is interpretable if it is unique and if all the documents are well represented by the topics. To perform a topic model of our dataset, it has been used the pyLDAvis python package.

3.2 Sentiment analysis

SA is the interpretation and classification of emotions within text data that can help to immediately identify critical situations, by processing efficiently huge amounts

of data, to take action right. It makes use of various Natural Language Processing (NLP) methods and algorithms, which may be based on:

- *Rule-based systems*, perform SA based on a series of manually processed instructions to help in the identification of subjectivity, polarity, or the subject of an opinion. To perform this step, these instructions incorporate different techniques, such as stemming, tokenization, part-of-speech tagging and parsing, or Lexicons (i.e. lists of words and expressions). A basic set of rules is made up of the following step: (1) Firstly, two lists of polarized words are defined (e.g. positive words such as generous, fabulous, happy, and negative words, such as fail, bad, sad); (2) Secondly, it needs counting how many positive or negative words occur in the specified text; (3) Thirdly, if the positive occurrences are greater than the negative ones, the text is flagged with a positive sentiment, and vice versa; finally, the text is flagged as neutral if the positive and negative occurrences are even.
- *Automatic systems*, which adopt machine learning techniques to learn from data. A SA task is usually modelled as a classification problem, where the classifier is trained using text to return a category, for example: positive, negative, or neutral. Usually, this task uses statistical models such as Naïve Bayes, Logistic Regression, Support Vector Machines, or Neural Networks. In the training process, the model learns to associate the corresponding *tags* (i.e. positive, negative, neutral) to a particular text based on the dataset used for training. The feature extractor allocates the input text into a feature vector, using an approach such as bag-of-words or bag-of-n-grams, all taking into consideration the frequency of the terms. Finally, in the prediction process, the model is fed by the feature vectors for generating the predicted tags.
- *Hybrid systems*, merge both rule-based and automatic approaches.

It is worth noting that, for training and testing automatic systems, there is the need to have available a set of resources, such as large datasets of marked texts or lexical databases in which each word has its polarity value. The problem that should be faced is that these resources are often very limited or completely missing, especially for non-English languages (Bosco et al. 2014).

In this paper, it has been adopted *Sentix* (Sentiment Italian Lexicon, Basile and Nissim, 2013), an available lexical resource for Italian SA based on the SVM classifier with a linear kernel and resulting from the alignment of several lexical and affective resources, such as *WordNet*—a large lexical database of English, *Multi-WordNet*—a multilingual lexical database (Pianta et al. 2002), *BabelNet*—a very large multilingual ontology (Navigli and Ponzetto 2012) and *SentiWordNet*—a lexical resource for opinion mining (Baccianella et al. 2010), in which 59,742 lemmas have been annotated for their polarity—whose scores range from -1 (totally negative) to 1 (totally positive), and intensity—whose scores range from 0 (totally neutral) to 1 (totally polarized), among other information. These quantities are used to derive $C_{\text{score}} = \text{Intensity} \times \text{Polarity}$ for each lemma, a score useful to divide lemmas into five groups: strongly positive: $0.25 \leq C_{\text{score}} < 1$; weakly positive: $0.125 \leq C_{\text{score}} < 0.25$; neutrals: $-0.125 \leq C_{\text{score}} \leq 0.125$; weakly negative: $-0.25 \leq C_{\text{score}} < -0.125$; strongly negative: $-1 \leq C_{\text{score}} < -0.25$. Moreover, since *Sentix* relies on *WordNet*

sense distinctions, one lemma could be associated with more than one C_{score} , and Sentix can treat them automatically. To this aim, it has been used *sentixR*, an R package useful to perform SA on Italian using Sentix, to extract a sentiment score for each tweet.

4 Results

In this paragraph, we present the analysis of the public debate during the *Italian Second Phase* of the pandemic, and the results obtained. To this aim, it has been chosen Twitter as the data source, as the user-generated contents are generally made up of Digital Context Data and, as aforementioned, this kind of data allows us to obtain insights from the public debate. Tweets are considered unstructured data, as they do not have a predefined model, nor are organized in a predefined manner. The period covered by our analysis goes from April 13th to June 15th, the day of transition to the *Italian Third Phase* in which the bathing establishments have been reopened and some restrictions have been abolished. Once the data source for the *Second Phase* monitoring has been decided, it has been necessary to understand how to search for those data that specifically speak only about tourism and Covid-19. One of the main issues is to direct the query and extract only the contents of our interest, being careful not to consider irrelevant data, and, on the other hand, not to restrict the search too much. Therefore, the combination of the search words clarified in Sect. 3 has been meticulously chosen to extract data.

During the monitoring phase, 3559 tweets has been collected. As aforementioned, special characters and links have been removed from texts, to easily analyse them. In Fig. 2, the trend of downloaded tweets during the observational period is shown.

As it can be seen, in the first weeks the number of gathered tweets was greater than in the last days, and perhaps this phenomenon is due to the fact that during the lockdown and at the beginning of the *Second Phase* troubles and uncertainties were more conspicuous than around the approach of the *Third Phase*. In the next Paragraphs, the analysis performed on these data are explained and discussions are done.

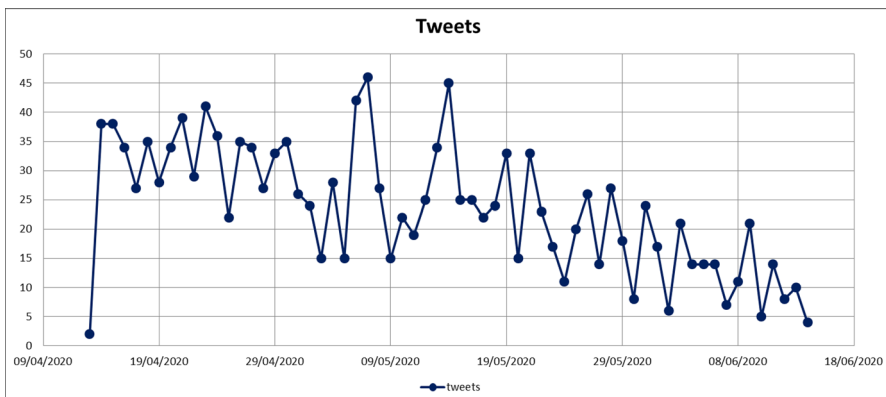


Fig. 2 Trend distribution of tweets during the observational period

4.1 LDA

A topic extraction through the LDA method has been performed on the dataset to try to extract the key topics during the discussions in the Covid-19 *Italian Second Phase*. In particular, from this dataset, all the re-tweeted posts have been removed as to not have duplicates during this analysis. As above described, there is not a given number of topics to be extracted from data, and different tests should be done to understand when each topic is interpretable, unique and if all the extracted topics are exhaustive for the whole dataset.

For these reasons, after several attempts, it has been fixed the number of clusters to be extracted to ten. In Fig. 3 the topic overview obtained by LDAvis is shown.

The left panel of this visualization presents a global view of the topic model and provide support to understand how prevalent is each topic and how topics are related to each other. Moreover, each topic is represented by a circle, whose centre is determined by calculating the distance between topics and using multidimensional scaling for projecting the inter topic distances onto two dimensions. Finally, the area of each circle is a measure of the prevalence of each topic within the dataset. The name of both axis, PC1 and PC2, refers to the fact that the algorithm used Principal Component as the algorithm for scaling the set of inter topic distances. As it can be seen, Topic 1, 2 and 8 are very distant from each other, while the others show an overlapping area of the circles. Thus, we expect that the terms representing clusters 1, 2 and 8, are quite different, consequently the meanings of each topic. The right panel of this visualization depicts a horizontal bar chart, and each bar represents the most salient terms of the whole dataset, determined by using *saliency*. These terms are useful to understand what is the meaning of each topic. The left and right panels together allow us to select each topic and to determine what are the most relevant

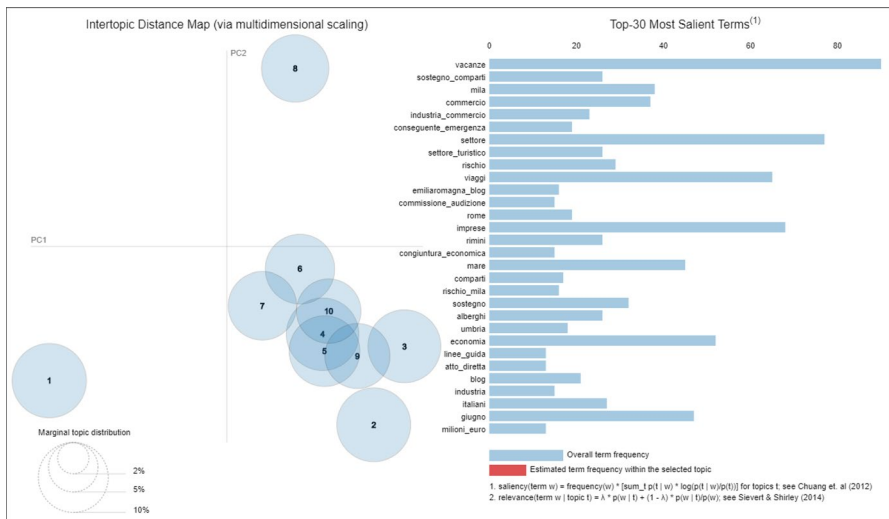


Fig. 3 The most salient terms—Topic Overview

terms of each cluster. When doing this process, the topic-specific frequency of each term is shown with respect to its corpus-wide frequency, and they are ranked by *relevance* for flexibly interpreting topics. Finally, as aforementioned, the parameter λ allows to better inspect the different contributions of each term within the topic. In Table 1 the most relevant terms for each cluster have been reported, with the final interpretation of the associated topics.

As it can be seen, Cluster 1, 2 and 8 show quite different topics: *Interventions to support the tourism sector* for Cluster 1, *Proposing solutions for restarting tourism* for Cluster 2 and *Building brand values of a destination* for Cluster 8. Moreover, many clusters have some traits in common, but these are associated with different aspects: for example, they often talk about the promotion of destinations but, while in Cluster 4 the topic is *Promoting an uplifting message of a destination to reduce crisis effects*, in Cluster 5 this theme is in relation to *health and safety of tourists*; in Cluster 9, they talk about the visibility of destinations through the *Promotion of the Italian cultural heritage to restart tourism*, while Cluster 10 deals with the issue of beaches, the reopening in security and the hospitality of a destination, thus it speaks about *New skills to reopen beaches and tourist facilities in total safety and tourist hospitality*. Thus, clusters show that different topics are interrelated to others but the combination of different arguments makes them unique. Finally, some topics are more sectorial and deal with specific issues: Cluster 3 is about *Crisis and social distancing in the restaurant sector*, while Cluster 7 addresses one of the fundamental problems inherent in tourism and concerns *Tourism emergency and travel restrictions*, as well as Cluster 6, which deals with the same issue but from the point of view of travellers' health, thus *Dealing with travel restrictions and safely travelling*.

Finally, in Fig. 4 the distribution of the tweets for each cluster is shown, while in Fig. 5 the daily trend per topic is shown.

It is quite evident that all the tweets are uniformly distributed among the different clusters and there is not a particular topic that emerges more than another one.

As for the daily trend of the topic, the daily discussion almost always embraces all topics, so we can say that there has never been a topic that has been discussed especially in the beginning, during or at the end of the *Second Phase*. This consolidates the fact that this phase was experienced with uncertainty in the future, trying to identify what could have been the possible or most probable solutions to restart the tourism sector, despite the national and international health situation due to the Covid-19 that allowed no medium/long-term projects.

4.2 Sentiment analysis

SA has been performed on the dataset to understand which is the sentiment emerged in the overall debate. As described above, *sentixR* has been used to accomplish this task. In Fig. 6, the sentiment trend in the monitored period is shown: green/red continuous lines indicate strongly positive/negative sentiment, green/red dashed lines indicate weakly positive/negative sentiment, and yellow line indicate neutral sentiment.

Table 1 The most relevant terms and the extracted topics for each cluster identified with LDA process

Clusters	Topic keywords	Topics
Cluster 1	tourism_sector, save_summer_season, help_enterprises, support_measures, retail, wine, formation, madeinitaly, Lazio_region	Topic1: Interventions to support the tourism sector
Cluster 2	proposed_solutions, opening_bathing_establishments, relaunch_decree, billion_euros, government_commitment, reopen, return_to_travel, abruzzo, cilento	Topic2: Proposing solutions for restarting tourism
Cluster 3	social_distancing, stay_home, restaurant_traders, we_accept_agreement, rome, umbria, trasimenolake, veneto, lazio, crisis, risk, work, enterprises, economy, season	Topic3: Crisis and social distancing in the restaurant sector
Cluster 4	art_exhibition, accommodation, summer_season, sicily, toscana, smartabruzzo, abbruzzoatyourhome, umbria, economic_revival, masks, support_measures	Topic4: Promoting an uplifting message of a destination to reduce crisis effects
Cluster 5	post_emergency, health_passport, study_safe_area, sardegna, tunisia, spain_reopens	Topic5: Promotion of destinations in relation to health and safety of tourists
Cluster 6	post_pandemic, guidelines, restart, travel_after, flows, travel_safe, planes, ryanair, italian_relaunch, womantimes_holiday, greece_reopens, veneto, egypt	Topic6: Dealing with travel restrictions and safely travelling
Cluster 7	tourism_sector_suffered_damage, lazio_regio_emergency, tourism_sector_affected, movement_between_regions, airplane_transport, l'aquila, veneto, germany, switzerland, campania	Topic7: Tourism emergency and travel restrictions
Cluster 8	relaunch_tourism_after_crisis, tour_operator, travel_agencies, sea, museums, carsulae_waterfall, destinations	Topic8: Building brand values of a destination
Cluster 9	culture, cultural_heritage, relaunch, promotion_plan, voucher, security, flor-ence, abruzzo	Topic9: Promotion of the Italian cultural heritage to restart tourism
Cluster 10	beach_umbrellas, new_competencies, ready_to_welcome, ready_to_reopen, holiday_bond, austria, spain, rimini, trentino, sardegna, german_tourists	Topic10: New skills to reopen beaches and tourist facilities in total safety and tourist hospitality

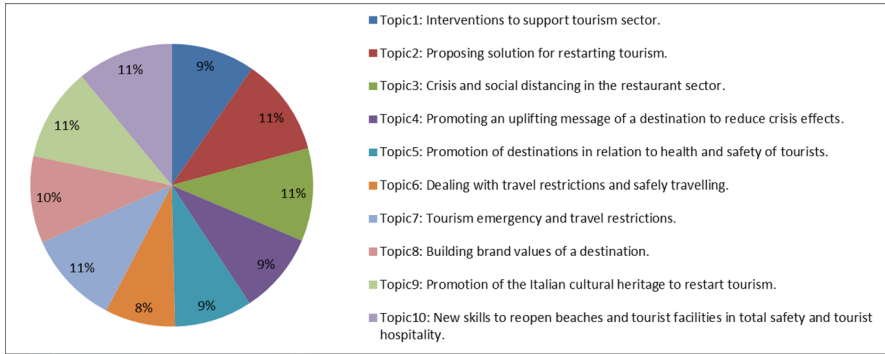


Fig. 4 Topic distribution

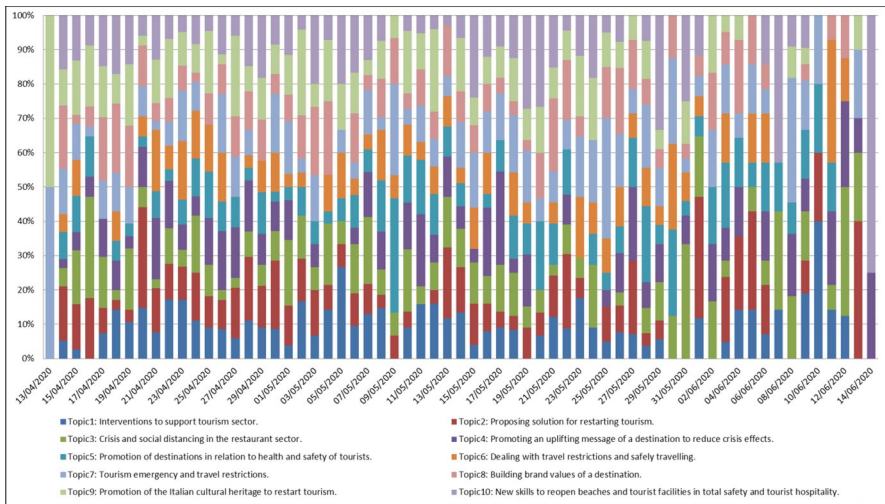


Fig. 5 Daily trend per topic

As it can be seen, the dominant sentiment of the whole debate is neutral during the whole period. Moreover, while at the beginning of the *Second Phase* weakly positive tweets were greater than weakly negative ones, during the last days this number slowed down. Furthermore, it is worth noting that there has never been a spike in strongly negative elements and that even weakly negative ones have always been below the positive threshold.

Besides this trend, sentiment analysis has been performed on every single topic. In Fig. 7 results are shown.

For all the topics, neutral sentiment is dominant on both the positive and negative slices, following the global trend just described. This means that, while speaking about a particular topic, there is not a strong component of positive or negative sentiment, which are a manifestation of, respectively, a favourable debate or an exacerbated

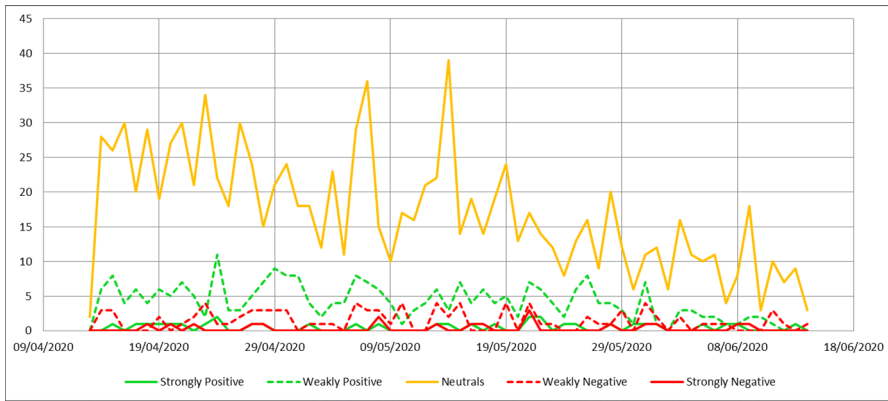


Fig. 6 Sentiment trend

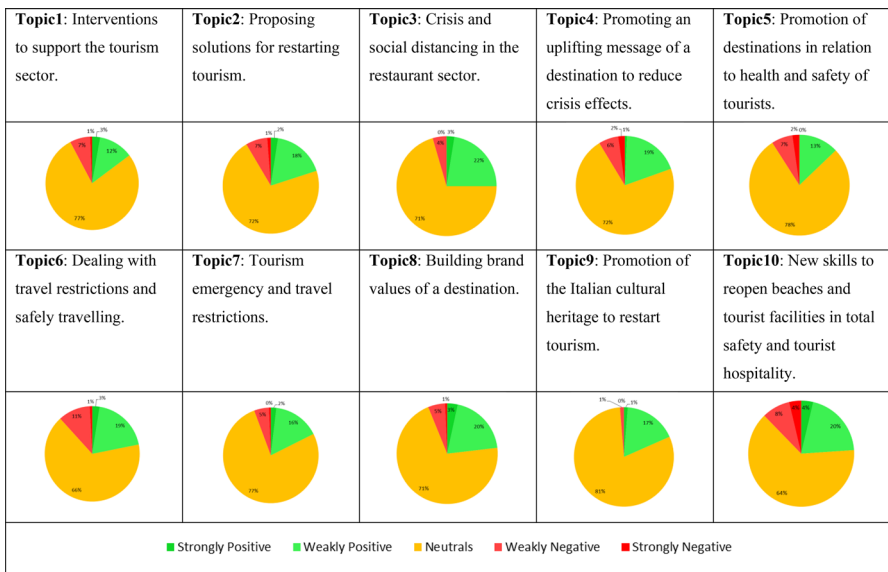


Fig. 7 Sentiment per topic

controversy. Instead, the debate is quite moderate, as they are probably trying to understand how the crisis could evolve and what kind of measures and precautions could be adopted to reduce impacts and find out feasible solutions. Moreover, Topic 9 shows the smallest part of both weakly and strongly negative sentiment, and this is probably because they are discussing how to promote destinations and how to associate good values to destinations brands. Despite, when the topics related to promotion and visibility of destinations are paired with others related to the safety and health of tourists and travel restrictions, the negative sentiment shows a greater area. Therefore, it is clear that these topics are the ones that have characterized the most negative comments and

concerns. On the contrary, the prevalence of neutral sentiment is a direct consequence of the period of uncertainty in which we were living, marked by the lack of decisions for tourism and health uncertainty at both the national and international levels.

5 Discussions

This paper demonstrates how the exploitation of Big Social Data and Analytics through the monitoring of virtual discussions aims at giving insights to businesses, governments and tourism operators for addressing the restart of tourism. For these purposes, data from Twitter has been collected in the *Italian Second Phase*, from April 13th to June 15th, the day of transition to the *Italian Third Phase* in which the bathing establishments have been reopened and some restrictions have been abolished. Then, these data have been cleaned up from re-tweeted posts and analysed by performing multi-modal analysis on the same dataset: topic extraction, with LDAvis (Sievert and Shirley 2014; Chuang et al. 2012) for their identification, prevalence and relation within the whole debate, and sentiment analysis, performed with SentixR (Sentiment Italian Lexicon—Basile and Nissim 2013) for the interpretation and classification of emotions within topics. In this section, findings from the analysis are presented and theoretical and practical contributions, as well as implications of the research, are discussed.

5.1 Evidence of the results

A preliminary analysis has shown that discussion has been more conspicuous at the beginning of the *Second Phase*, probably because troubles and uncertainties were greater than those during the approach to the *Third Phase*.

Results from LDA have shown that several topics have been treated in the online social debate. In particular, *Interventions to support the tourism sector*, *Proposing solutions for restarting tourism* and *Building brand values of a destination* characterise the main discussion in three topics that are distant among them. Besides, another important topic is related to *Crisis and social distancing in the restaurant sector*, as social distancing measures have heavily impacted the foodservice industry, on one hand in rethinking strategies for guaranteeing hospitality and, on the other hand, in redesigning workplaces to fit in with social distancing requirements.

Moreover, the combination of several keywords belonging to different contents within the same cluster makes each topic unique during the debate. As for *Promotion of destination*, it can be noted that it is necessary to boost *an uplifting message of a destination to reduce crisis effects*, to exploit natural beauty, landscape and *the Italian cultural heritage to restart tourism* and to develop *New skills to reopen beaches and tourist facilities in total safety and tourist hospitality*, and these aspects must be taken into consideration when dealing with tourism restarting. At the same time, the analysis has shown that, when promoting destinations, some perplexities and uncertainties have arisen when dealing with the *health and safety of tourists* as

well as *travel restrictions*, the latter linked both to the emergency which tourism is facing up to and to the measures that should be taken to *safely travel*.

Furthermore, as shown in the previous sections, the daily discussion has almost always embraced all topics, thus there has never been a topic that has been discussed especially in the beginning, during or at the end of the *Second Phase*. This consolidates the fact that this phase was experienced with uncertainty in the future, trying to identify what could have been the possible or most probable solutions to restart the tourism sector, despite the national and international health situation due to the Covid-19 that allowed no medium/long-term projects.

Another result in support of these considerations has been given by sentiment analysis, which has shown that the dominant sentiment is neutral during the whole period and within each topic. Thus, the prevalence of neutral sentiment is a direct consequence of the period of uncertainty in which Italy has been, marked by the lack of decisions for tourism and health uncertainty at both national and international levels. Starting from these shreds of evidence, the decision-makers could intervene giving greater guarantees to the sector and foreseeing and/or providing more concrete actions to help the sector to minimize the effects of the crisis and encourage tourism. What emerges from the analysis is that many tourist destinations have tried to increase their visibility and promote them, trying to give an uplifting message and to associate positive values with their brand.

Finally, it has emerged that if on the one hand the tour operators and the different stakeholders have tried to promote their destinations, on the other hand, a request for help addressed to the institutions has emerged to obtain concrete help from them to face the crisis that was going through, for guaranteeing jobs safeguarding and being able to save their industrial reality.

Through a more accurate analysis of the topics and of the keywords that characterise each of them, it has been possible to extract additional information, with the final aim of improving the tourist offer and recovering the tourism sector. The extracted topics can be grouped into three macro discussion areas: *Industry sustainable policies and targeted support* (Topic 1 and 2), *Destination promotion and brand management* (Topic 4, 5, 8, 9 and 10), *Travel, safety and social restrictions* (Topic 3, 6 and 7).

As for the *Industry sustainable policies and targeted support* macro discussion area, related issues coming from tourist stakeholders comprises requests for both short and mid-term government support. In fact, the debate highlighted the opportunity to safeguard the upcoming summer tourist season and to sustain retails with a particular focus on “Made in Italy” productions. At the same time, stakeholders were looking for mid-term sustainable solutions to reopen and maintain bathing establishments or to return to travel safely. The second macro discussion area, *Promotion of destination*, highlights some keywords that are connected to the acquisition of new competencies to reopen and to be ready to welcome tourists, the promotion of new services connected to smart services (i.e. Smart Abruzzo or “Abruzzo at your home”) or specific local natural resources to make appealing messages and building a positive brand associated to that destination. These examples could be good practises for other destinations as they could provide them with new ideas, and act as an engine to identify common challenges and create territorial synergies or

stakeholders' networks and partnerships, to improve the tourist offer of the tourism destinations. The third macro discussion area that emerged in our analysis, *Travel, safety and social restrictions*, helps to understand tourists and locals' needs. As also reported in the UNWTO (2020b), 100% of global destinations continued to adopt travel restrictions, while 72% of these destinations (including Italy) had their borders completely closed. A gradual responsible reopening started in June has made it possible to save jobs and allow tourism to once again take on its vital role in driving sustainable development. On the other hand, as it emerges from keywords, social distancing has put a strain on the restaurateurs. They have always looked for solutions to guarantee conviviality and hospitality in compliance with the social distancing restrictions, ready for any kind of agreement with policymakers and stakeholders (i.e. "we_accept_agreement"). Moreover, in the process of reorganization for restarting their own business, all the tourism stakeholders have deeply examined how to create a "safe area" (i.e. "study_safe_area") for tourists and local people, to guarantee all the specific requirements, especially in the bathing establishments.

5.2 Research contributions and implications

This paper aims at contributing to the literature by providing a methodological framework, based on Big Data and Analytics, which helps tourist destinations to improve the speed and effectiveness of the response to a crisis, reducing post-crisis recovery issues for a STD. In particular, this paper shows how Big Data and Analytics can help in fastening the process of crisis rationalisation by hindsight when unpredictable critical events occur (Aldao et al. 2021), providing real-time answers to STD stakeholders and decision-makers on issues like the damage to destination image and reputation, and the changes in tourist behaviour following crises and disasters (Mair et al. 2016). The presented methodological framework is a potential and powerful decision-making tool, as it allows to identify what could be the actions that should be taken by governments or businesses to tackle the restart of the tourism sector through the monitoring of topics and sentiment expressed by people on social media. These insights are useful and can be the basis for the development of adequate policy responses. This will ensure policymakers keep a broad view during their support interventions and government to assume a participatory role for enhancing sustainable development and economic growth in this sector, a fundamental step for becoming a STD (Gretzel et al. 2015; Caragliu et al. 2011).

The analytics implemented are based on the BSD multimodal analytics approach (Solazzo et al. 2020, 2021b). The application of this approach to the debate on the Covid-19 pandemic for evaluating the impact on the tourism sector represents an original contribution. Indeed, the combination of topics extraction and sentiment analysis allows us to identify at the same time which are the "hot topics" and the feelings that people are talking with about them. From this perspective, the presented analysis provides a further contribution to the advancement of the research on digital tourism and social media platforms for destination management. Specifically, the study confirms that social media platforms like Twitter are, indeed, recognized as the social network more focused on textual content and perceived as a

conversational platform oriented to the present and its immediate development (MacKay et al. 2017). Moreover, this study has demonstrated that tourism stakeholders can exploit social media to grab particular ideas, which arise from the problems and needs highlighted from the wisdom of crowds, and which can act as an incentive to create a new business or to accelerate the process for creating a STD (Buhalis and Amaranggana 2013).

In contributing to the research on STDs, the paper provides evidence about the role that digital technologies can assume in the strategies for the restarting of tourism companies as well as for governments involved in the planning and execution of public policies for sustaining the competitiveness of tourism companies and destinations and managing the crisis. Digital technologies are assumed to be crucial in the rebuilding of tourism that, as highlighted by the OECD (2020), has to become more sustainable and resilient. The expected process of transformation needs to be supported by political actions able to reduce the losses and to promote a coordinated strategy able to reduce the negative impacts of the emergence by creating opportunities for future development. According to the OECD (2020), the priorities of the re-starting have to look at the restoration of confidence of travellers, to reduce the uncertainty with clear information and coordinated political initiatives also at countries' level. In all these actions, depicting the profile of a STD, digital technologies are fundamental features for assuring the transition of tourism towards an information-intensive configuration. The responsibility of governments and public agencies is relevant also in the communication strategies undertaken to inform citizens and tourists. For this purpose, the re-starting requires a well-conceived communicational and marketing program that has to be able to reduce uncertainty by mixing messages of assurance and interest for a safe but pleasant experience (Ketter and Avraham 2021).

This paper also demonstrates how a stronger focus on smartness and real-time can condense space and time and increase a destination's adaptive response capacity (Bethune et al. 2022). Indeed, the real-time analysis of these data, which is a fundamental component for STD development (Buhalis and Amaranggana 2015), can contribute within the presented methodology to develop the emerging body of work on Destination Resilience and in particular on the integration of resilience thinking into operational and decision-making settings of a STD.

Finally, this work suggests that becoming a STD facilitates an integrated approach (Moe and Pathranarakul 2006) to crisis management in which both proactive and reactive strategies enable the stakeholders to respond to the crisis, before, during, and after the occurrence. Both proactive and reactive strategies are normally required to minimise chaos and social disorders during a time of crisis (Antony and Jacob 2019). The proposed methodology could be part of an integrated approach to crisis management, as the insights produced can be used by a STD either to drive mitigations, preparedness and warnings in proactive strategies or to model the reactive strategies after the events occurred.

This study offers also some practical contributions. Through this methodology, it is possible to identify insights for short and mid-term government support as emerged from *Industry sustainable policies and targeted support* macro discussion area. Furthermore, as emerged from the macro discussion area *Promotion of*

destination, it is possible to highlight which are the more proactive regions as self-tourist promoters and if, possibly, they are adopting smart technologies (such as virtual and augmented reality) or strategic measures to be both as smart and hospitable as possible in the pandemic and to reduce the impacts of restrictions relating to containment of the infection from Covid-19. Moreover, the methodology provides a powerful tool to understand tourists and locals' needs through the exploitation of BSD as emerged from the macro discussion area of *Travel, safety and social restrictions*.

Implications of this work are twofold: for researchers, this approach could be used along with other methods of data collection and analysis, to obtain more complete insights. For practitioners, this work could be a strategic tool as it provides important information useful for monitoring the impact of a crisis event on the performance of the tourism sector, and for offering valuable insights to governments, destination managers and hotel managers for the development of proper policy actions. In particular, by examining the extracted topics, it could be possible to understand how tourism stakeholders are aware of the situation and able to proactively react to the pandemic, or it could suggest interventions and more strategies for managing the crisis and helping to rebuild businesses affected by this pandemic event.

This study allows also to derive practical implications for tourism companies by confirming the contribution that digitalization can provide at the conception and implementation of data-driven business models, the competitive positioning on niches market with higher margins of profit, a more effective strategy of segmentation, the creation of a personalized and tailor-made offering of tourism experiences, and the smart configuration of the destinations (Pasquinelli and Trunfio 2020).

Therefore, the continuous monitoring of these discussions on social media allows to exploit the wisdom of crowds and to have specific insights on the expectations of both tourists, tour operators, tourism stakeholders in general and local people of the tourist destinations, thus providing an overall view, allowing choices to be made in keeping with the needs came out and giving the possibility to share good practices between them. Among the six future research paths identified in Zenker and Kock (2020) as a starting point for a research agenda, this paper aims at giving a practical contribution to understand how Big Data Analytics and Machine Learning allow extracting useful insights from BSD in the tourism sector, to improve aspects such as changes in Destination Image and changes in the Tourism Industry.

6 Conclusions

The Covid-19 pandemic is impacting several aspects of our lives and, among the different sectors, tourism is being hugely affected (UNWTO 2020c). In this new context, tourism managers and stakeholders need to look ahead to the future, trying to wisely react to overcome the crisis in this sector.

This paper presents an application of the BSD paradigm and Machine Learning techniques for monitoring the public debate on social media, to obtain deeper insights about Covid-19 effects on the tourism sector, offering valuable insights on how tourism stakeholders and citizens are talking about the crisis. In particular, our

study suggests that the STD framework is a fundamental paradigm when unforeseen challenges, from a management point of view, arise to handle crises and disasters promptly.

The limitations of this paper are due to the fact that only one data source (Twitter) has been used to perform the analysis: other Online Social Networks, such as Facebook, can be used for improving the number of comments and discussions and the quality of the performed analysis. Moreover, the application of Network Analysis and metrics can help to extract insights about the network and to understand how the discussion is spread through the social network or if the smallest communities speak about some particular topics than others.

Future research can adopt the same approach for performing real-time analysis and monitoring e-WOM to investigate changes in Tourism Behaviours and changes in Residents Behaviours, two research paths identified in Zenker and Kock (2020), which are both important when dealing with hospitality within a STD. Moreover, this analysis could be further enriched with Social Network Analysis and the profiling of all the users who intervene within the debate, providing more details and a better perspective when studying the support interventions to be implemented. Finally, this study has empirically tested the framework during the *Italian Second Phase* of the Covid-19 pandemic, but its usage could be extended in other contexts.

References

- Abdar M, Zomorodi-Moghadam M, Das R, Ting IH (2017) Performance analysis of classification algorithms on early detection of liver disease. *Expert Syst Appl* 67:239–251
- Aldao C, Blasco D, Espallargas MP, Rubio SP (2021) Modelling the crisis management and impacts of 21st century disruptive events in tourism: the case of the COVID-19 pandemic. *Tour Rev*
- Aldunce P, Beilin R, Howden M, Handmer J (2015) Resilience for disaster risk management in a changing climate: practitioners' frames and practices. *Glob Environ Change* 30:1–11
- Antony JK, Jacob JM (2019) Crisis management in the tourism industry—the role of social media platforms. *Atna J Tour Stud* 14(1):83–96
- Ardito L, Cerchione R, Del Vecchio P, Raguseo E (2019) Big data in smart tourism: challenges, issues and opportunities. *Curr Issues Tour* 22(15):1805–1809
- Avis M (2003) Do we need methodological theory to do qualitative research? *Qual Health Res* 13(7):995–1004
- Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. *Lrec* 10:2200–2204
- Basile V, Nissim M (2013) Sentiment analysis on Italian tweets. In: *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pp. 100–107
- Bethune E, Buhalis D, Miles L (2022) Real time response (RTR): conceptualizing a smart systems approach to destination resilience. *J Dest Mark Manage* 23:100687
- Bosco C, Allisio L, Mussa V, Patti V, Ruffo G F, Sanguinetti M, Sulis E (2014) Detecting happiness in Italian tweets: towards an evaluation dataset for sentiment analysis in Felicità. In: *5th international workshop on emotion, social signals, sentiment & linked open data, ES³LOD 2014 European Language Resources Association*, pp 56–63
- Brandt T, Bendler J, Neumann D (2017) Social media analytics and value creation in urban smart tourism ecosystems. *Inf Manag* 54(6):703–713
- Buhalis D (2019) Technology in tourism—from information communication technologies to eTourism and smart tourism towards ambient intelligence tourism: a perspective article. *Tour Rev*

- Buhalis D, Amaranggana A (2013) Smart tourism destinations. Information and communication technologies in tourism 2014. Springer, Cham, pp 553–564
- Buhalis D, Amaranggana A (2015) Smart tourism destinations enhancing tourism experience through personalisation of services. Information and communication technologies in tourism 2015. Springer, Cham, pp 377–389
- Buhalis D, Foerste M (2015) SoCoMo marketing for travel and tourism: empowering co-creation of value. *J Dest Mark Manage* 4(3):151–161
- Bulchand-Gidumal J (2022) Post-COVID-19 recovery of island tourism using a smart tourism destination framework. *J Dest Mark Manage* 23:100689
- Caragliu A, Del Bo C, Nijkamp P (2011) Smart cities in Europe. *J Urban Technol* 18(2):65–82
- Caragliu A, Del Bo C, Nijkamp P (2013) Smart cities in Europe. VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics, Research Memoranda 0048, Amsterdam, The Netherlands, 2009 From Moscow to Sao Paulo, Emerging 7 Cities Report
- Carter SM, Little M (2007) Justifying knowledge, justifying method, taking action: epistemologies, methodologies, and methods in qualitative research. *Qual Health Res* 17(10):1316–1328
- Cavlek N (2002) Tour operators and destination safety. *Ann Tour Res* 29(2):478–496
- Chuang J, Manning C D, Heer J (2012) Termite: visualization techniques for assessing textual topic models. In: Proceedings of the international working conference on advanced visual interfaces, pp 74–77
- Chun J, Kim C K, Kim G S, Jeon J, Lee W K (2020) Social big data informs spatially explicit management options for national parks with high tourism pressures. *Tour Manag.* 81: 104136.
- Davenport TH (2013) Turning towards a smarter travel experience. Amadeus IT Group. 17
- Del Chiappa G, Baggio R (2015) Knowledge transfer in smart tourism destinations: analyzing the effects of a network structure. *J Dest Mark Manage* 4(3):145–150
- Del Vecchio P, 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
- Desdemoustier J, Crutzen N, Cools M, Teller J (2019) Smart City appropriation by local actors: an instrument in the making. *Cities* 92(September):175–186
- Duran G, Uygur SM (2022) A comprehensive systematic literature review about smartness in tourism. Handbook of research on digital communications, Internet of Things, and the Future of Cultural Tourism, pp 203–227
- Faulkner B (2001) Towards a framework for tourism disaster management. *Tour Manag* 22(2):135–147
- Femenia-Serra F, Ivars-Baidal JA (2018) Do smart tourism destinations really work? The case of Benidorm. *Asia Pacific J Tour Res* 1–20.
- Femenia-Serra F, Neuhofer B, Ivars-Baidal JA (2019) Towards a conceptualisation of smart tourists and their role within the smart destination scenario. *Serv Ind J* 39(2):109–133
- Filieri R (2016) What makes an online consumer review trustworthy? *Ann Tour Res* 58:46–64
- Forbes (2020) How virtual reality could help the travel & tourism industry in the aftermath of the coronavirus outbreak. www.forbes.com/sites/solrogers/2020/03/18/virtual-reality-and-tourism-whats-already-happening-is-it-the-future/#1395e8b28a6a . Accessed Feb 2022
- Fuchs M, Höpken W (2011) E-Business horizons in the tourism industry-Challenges for research and practice. In Food, Agri-Culture and Tourism. Springer, Berlin, pp. 140–160.
- García-Milon A, Juaneda-Ayensa E, Olarte-Pascual C, Pelegrín-Borondo J (2020) Towards the smart tourism destination: key factors in information source use on the tourist shopping journey. *Tour Manag Perspect* 36:100–730
- Gössling S, Scott D, Hall C M (2020) Pandemics, tourism and global change: a rapid assessment of COVID-19. *J Sustain Tour.* 1–20.
- Gretzel U, Scarpino-Johns M (2018) Destination resilience and smart tourism destinations. *Tour Rev Int* 22(3–4):263–276
- Gretzel U, Yuan YL, Fesenmaier DR (2000) Preparing for the new economy: advertising strategies and change in destination marketing organizations. *J Travel Res* 39(2):146–156
- Gretzel U, Sigala M, Xiang Z, Koo C (2015) Smart tourism: foundations and developments. *Electron Mark* 25(3):179–188
- Gretzel U, Zhong L, Koo C (2016) Application of smart tourism to cities. *Int J Tourism Cities* 2:2
- Gretzel U, Ham J, Koo C (2018) In: Wang Y, Shakeela A, Kwek A, Khoo-Lattimore C (eds) Managing Asian destinations. Perspectives on Asian tourism. Springer, Singapore
- Hajibaba H, Karlsson L, Dolnicar S (2017) Residents open their homes to tourists when disaster strikes. *J Travel Res* 56(8):1065–1078

- Hall C M, Prayag G, Amore A (2017) Tourism and resilience: individual, organisational and destination perspectives. Channel View Publications
- Hesse-Biber SN, Leavy P (2011) Focus group interviews. The practice of qualitative research, pp 163–192
- Höpken W, Fuchs M, Lexhagen M (2019) Big data analytics for tourism destinations. In: Advanced methodologies and technologies in network architecture, mobile computing, and data analytics. IGI Global, pp. 28–45
- Hu YH, Chen YL, Chou HL (2017) Opinion mining from online hotel reviews: a text summarization approach. *Inf Process Manag* 53(2):436–449
- Iacus SM, Natale F, Santamaria C, Spyros S, Vespe M (2020) Estimating and projecting air passenger traffic during the COVID-19 coronavirus outbreak and its socio-economic impact. *Saf Sci* 129:104–791
- Iorio C, Pandolfo G, D'Ambrosio A, Siciliano R (2020) Mining big data in tourism. *Qual Quant* 54(5):1655–1669
- Jia Z, Shi Y, Jia Y, Li D (2012) A framework of knowledge management systems for tourism crisis management. *Proc Eng* 29:138–143
- Jovicic DZ (2019) From the traditional understanding of tourism destination to the smart tourism destination. *Curr Issue Tour* 22(3):276–282
- Ketter E, Avraham E (2021) # StayHome today so we can# TravelTomorrow: tourism destinations' digital marketing strategies during the Covid-19 pandemic. *J Travel Tour Mark* 38:1–14
- Kim H, Joun HJ, Choe Y, Schroeder A (2019) How can a destination better manage its offering to visitors? Observing visitor experiences via online reviews. *Sustain* 11(17):4660
- Kirk CP, Rifkin LS (2020) I'll trade you diamonds for toilet paper: consumer reacting, coping and adapting behaviors in the COVID-19 pandemic. *J Bus Res* 11:124–131
- Kushwaha AK, Kar AK, Dwivedi YK (2021) Applications of big data in emerging management disciplines: a literature review using text mining. *Int J Inf Manage Data Insights* 1(2):1000–17
- Lalicic L, Önder I (2018) Residents' involvement in urban tourism planning: opportunities from a smart city perspective. *Sustainability* 10(6):1852
- Li J, Xu L, Tang L, Wang S, Li L (2018) Big data in tourism research: a literature review. *Tour Manage* 68:301–323
- Liamputtong P, Ezzy D (2005) Qualitative research methods. Second. Oxford University Press, Melbourne
- MacEachren AM, Cai G (2006) Supporting group work in crisis management: visually mediated human–GIS–human dialogue. *Environ Plann B Plann Des* 33:435–456
- MacKay K, Barbe D, Van Winkle CM, Halpenny E (2017) Social media activity in a festival context: temporal and content analysis. *Int J Contemp Hosp Manag* 29(2):669–689
- Mair J, Ritchie BW, Walters G (2016) Towards a research agenda for post-disaster and post-crisis recovery strategies for tourist destinations: a narrative review. *Curr Issues Tour* 19(1):1–26
- Malizia A, Bellucci A, Diaz P, Aedo I, Levialdi S (2011) eStorys: a visual storyboard system supporting back-channel communication for emergencies. *J Vis Lang Comput* 22(2):150–169
- Marine-Roig E, Clavé SA (2015) Tourism analytics with massive user-generated content: a case study of Barcelona. *J Dest Mark Manage* 4(3):162–172
- Maruccia Y, Mele G, Ndou V, Del Vecchio P (2019) A methodological framework for measuring the smartness of Tourism Destinations. In: IFKAD 2019-international forum on knowledge asset dynamics, pp 92–107
- McKercher B, Chon K (2004) The over-reaction to SARS and the collapse of Asian tourism. *Ann Tour Res* 31(3):716
- Menk A, Sebastia L, Ferreira R (2017) Curumim: A serendipitous recommender system for tourism based on human curiosity. In: 2017 IEEE 29th international conference on tools with artificial intelligence (ICTAI). IEEE, pp 788–795
- Miah SJ, Vu HQ, Gammack J, McGrath M (2017) A big data analytics method for tourist behaviour analysis. *Inf Manag* 54(6):771–785
- Micera R, Presenza A, Del Chiappa G (2013) Smart destinations. New strategies to manage tourism industry. In: IFKAD 2013-international forum on knowledge asset dynamics, pp 1405–1422
- Moe TL, Pathranarakul P (2006) An integrated approach to natural disaster management: public project management and its critical success factors. *Disaster Prev Manag Int J*
- Montero AA, López-Sánchez JA (2021) Intersection of data science and smart destinations: a systematic review. *Front Psychol* 12

- Navigli R, Ponzetto SP (2012) BabelNet: the automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artif Intell* 193:217–250
- Nguyen TT, Camacho D, Jung JE (2017) Identifying and ranking cultural heritage resources on geotagged social media for smart cultural tourism services. *Pers Ubiquitous Comput* 21(2):267–279
- OECD (2020) “Policy Responses to Coronavirus (COVID-19) Rebuilding tourism for the future: COVID-19 policy responses and recovery”. <https://www.oecd.org/coronavirus/policy-responses/rebuilding-tourism-for-the-future-covid-19-policy-responses-and-recovery-bced9859/>
- Olshannikova E, Olsson T, Huhtamäki J, Kärkkäinen H (2017) Conceptualizing big social data. *J Big Data* 4(1):1–19
- Ordoñez de Pablos P, Turro LJ, Tennyson RD, Zhao J (2015) Knowledge management for competitive advantage during economic crisis. IGI Global, Hershey
- Park JH, Lee C, Yoo C, Nam Y (2016) An analysis of the utilization of Facebook by local Korean governments for tourism development and the network of smart tourism ecosystem. *Int J Inf Manage* 36(6):1320–1327
- Pasquinelli C, Trunfio M (2020) Overtouristified cities: an online news media narrative analysis. *J Sustain Tour* 28(11):1805–1824
- Pasquinelli C, Trunfio M (2021) The missing link between overtourism and post-pandemic tourism. Framing Twitter debate on the Italian tourism crisis. *J Place Manag Dev*
- Pianta E, Bentivogli L, Girardi C (2002) Developing an aligned multilingual database. In: *Proceedings of Global WordNet Conference*
- Prayag G (2018) Symbiotic relationship or not? Understanding resilience and crisis management in tourism. *Tour Manag Perspect* 25:133–135
- Raguseo E, Neirotti P, Paolucci E (2017) How small hotels can drive value their way in infomediation. The case of ‘Italian hotels vs. OTAs and TripAdvisor.’ *Inf Manage* 54(6):745–756
- Schroeder A, Pennington-Gray L, Donohoe H, Kiouis S (2013) Using social media in times of crisis. *J Travel Tour Mark* 30(1–2):126–143
- Sievert C, Shirley K (2014) LDAvis: A method for visualizing and interpreting topics. In: *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, pp 63–70
- Sigala M (2020) Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *J Bus Res* 117:312–321
- Sigala M, Marinidis D (2012) Web map services in tourism: a framework exploring the organizational transformations and implications on business operations and models. *Int J Bus Inf Syst* 9(4):415–434
- Sigala M, Christou E, Gretzel U (eds) (2012) *Social media in travel, tourism and hospitality: theory, practice and cases*. Ashgate Publishing, Ltd.
- Sigalat-Signes E, Calvo-Palomares R, Roig-Merino B, García-Adán I (2020) Transition towards a tourist innovation model: the smart tourism destination: reality or territorial marketing? *J Innov Knowl* 5(2):96–104
- Solazzo G, Maruccia Y, Lorenzo G, Elia G, Del Vecchio P, Ndou V (2020) Extracting Insights from Big Social Data for Smarter Tourism Destination Management. In *IFKAD 2020-Knowledge in Digital Age*. IFKAD.
- Solazzo G, Elia G, Passiante G (2021a) Defining the big social data paradigm through a systematic literature review approach. *J Knowl Manage*
- Solazzo G, Maruccia Y, Lorenzo G, Ndou V, Del Vecchio P, Elia G (2021b) Extracting insights from big social data for smarter tourism destination management. *Meas Bus Excell*.
- Steiner C, Richter T, Dörny S, Neisen V, Stephenson ML, Lemma AF, Mitchell JG (2012) Economic crisis, international tourism decline and its impact on the poor. *Economic crisis, international tourism decline and its impact on the poor*
- Sułkowski Ł (2020) Covid-19 pandemic; recession, virtual revolution leading to de-globalization? *J Intercult Manage* 12(1):1–11
- Sutton JN, Palen L, Shklovski I (2008) Backchannels on the front lines: Emergency uses of social media in the 2007 Southern California Wildfires. In: *Proceedings of the Conference on Information Systems for Crisis Response and Management (ISCRAM)*
- Taleb NN (2007) *The black swan: The impact of the highly improbable (vol 2)*. Random house
- Trunfio M, Pasquinelli C (2021) Smart technologies in the Covid-19 crisis: managing tourism flows and shaping visitors’ behaviour. *Eur J Tour Res* 29:2910–2910

- UNWTO (2020a) Supporting jobs and economies through travel & tourism—a call for action to mitigate the socio-economic impact of COVID-19 and accelerate recovery. The World Tourism Organization (UNWTO). <https://doi.org/10.18111/9789284421633>
- UNWTO (2020b) Travel restrictions 4th the World tourism Organization (UNWTO). <https://www.unwto.org/news/covid-19-restrictions-on-tourism-travel>. Accessed Feb 2022
- UNWTO (2020c) Impact assessment of the Covid-19 outbreak on international tourism. <https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism>. Accessed Aug 2020
- Vieweg S, Palen L, Liu S B, Hughes A L, Sutton J (2008) Collective intelligence in disaster: examination of the phenomenon in the aftermath of the 2007 Virginia Tech shootings. In: Proceedings of the 5th International ISCRAM Conference, Washington, DC, pp 44–54
- Wang D, Li X, Li Y (2013) China's "smart tourism destination" initiative: A taste of the service-dominant logic. *J Dest Mark Manage* 2(2):59–61
- World Economic Forum (2020) These new gadgets were designed to fight COVID-19. <https://www.weforum.org/agenda/2020/04/coronavirus-covid19-pandemic-gadgets-innovation-technology/>. Accessed Feb 2022
- Xiang Z, Schwartz Z, Gerdes JH Jr, Uysal M (2015a) What can big data and text analytics tell us about hotel guest experience and satisfaction? *Int J Hosp Manag* 44:120–130
- Xiang Z, Wang D, O'Leary JT, Fesenmaier DR (2015b) Adapting to the internet: trends in travelers' use of the web for trip planning. *J Travel Res* 54(4):511–527
- Xie G, Qian Y, Wang S (2021) Forecasting Chinese cruise tourism demand with big data: an optimized machine learning approach. *Tour Manag* 82:104208
- Yu M, Li Z, Yu Z, He J, Zhou J (2021) Communication related health crisis on social media: a case of COVID-19 outbreak. *Curr Issue Tour* 24(19):2699–2705
- Zenker S, Kock F (2020) The coronavirus pandemic: a critical discussion of a tourism research agenda. *Tour Manag* 81:104164

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