

Valerie Stopp

Destination Image Online
analyzed through
User Generated Content
A Systematic
Literature Review



University of Algarve
Faculty of Economics

2021

CD

MASTER THESIS

Destination Image Online
analyzed through
User Generated Content
A Systematic
Literature Review

Valerie Stopp

Master in Management, specialization: Tourism Management

Research made under the supervision of:

Professor Doutor Nelson Matos

Professora Doutora Célia Ramos

Work Authorship Declaration

Destination Image Online analyzed through User Generated Content.

A Systematic Literature Review

I declare to be the author of this work, which is unique and unprecedented. Authors and works consulted are properly cited in the text and are included in the listing of references.

Valerie Stopp

Frankfurt, 7/15/2021.....

Copyright

© Copyright: Valerie Stopp

The University of Algarve reserves the right, in accordance with the provisions of the Portuguese Copyright and Related Rights Code, to archive, reproduce and make public this work, regardless of means used, as well as to broadcast it through scientific repositories and allow its copy and distribution with merely educational or research purposes and non-commercial purposes, provided that credit is given to the respective author and publisher.

Dedication

Writing a master thesis is a learning process. One learns as much about the studied subject as about oneself. For me one of the biggest learnings in the years 2020 / 21, was that time on earth is limited. Hence, one sometimes must prioritize – to squeeze in all the wanted to do's – in the limited time.

I thank my family and friends – especially one very dear friend – for choosing to share their time with me. I appreciate it a lot.

A big thank you also goes to my two supervisors. For their patience and kind support.

Finally, I thank the University of Algarve for giving me the possibility of having a lovely experience abroad during my master studies.

Abstract

Destination Image is a concept that has been studied for a long time in tourism research. The question of how a destination is perceived by tourists and potential new guests is an important insight, especially for local tourism managers, in order to evaluate the implemented strategies and to plan further tactics.

Since the last two decades, due to a drastic digitalization, tourism research is now increasingly examining the Destination Image online. This creates new challenges in the selection of sources, methods, and in data collection.

The aim of the present study was to systematically capture the approach to analyze the online Destination Image through User Generated Content using studies from the last ten years. Therefore, a Systematic Literature Review on primary research from academic databases was conducted.

As a summary of the findings, a conceptual model was developed, based on the insights of the studies in the dataset, to contribute a guidance for the preparation phase of future online Destination Image research.

In short, the main findings are: TripAdvisor.com is the main source for online Destination Image analysis. Researchers recommend using the help of software and programming languages to collect and analyzed the data. Equally to earlier Destination Image studies, the main methods applied in online Destination Image analysis are quantitative content analysis, qualitative content analysis and sentiment analysis. In combination with the examination of cognitive and affective factors, co-occurrence analysis, and correlation analysis.

The present study has several limitations, which are: the loss of detail information due to reducing the studies to comparable key parameters, the absence of Anglo-American studies, due to the database selection as well as the lack of quality testing of the studies included.

Keywords: Online Destination Image, Systematic Literature Review, Tourism Research, Tourist Generated Content

Resumo

A *Destination Image* é um conceito que tem sido estudado há muito tempo na investigação turística. A questão de como o destino é visto pelos turistas e pelos potenciais novos hóspedes é uma perspectiva importante, especialmente para os gestores de turismo da região, a fim de avaliar as estratégias implementadas e de planear novas táticas.

Desde as últimas duas décadas, ocorreu uma digitalização drástica, a investigação turística adaptou-se a este fenómeno e está agora a estudar cada vez mais a imagem do destino online. Esta alteração criou novos desafios na selecção de fontes, métodos, e na recolha de dados.

O objetivo do presente trabalho foi o de captar, de forma sistemática, as abordagens consideradas para analisar a imagem do destino online utilizando estudos dos últimos dez anos. Para este efeito, os estudos primários dos anos 2010-2020 das bases de dados académicos Web of Science, ProQuest e b-on, foram recolhidos utilizando palavras-chave de pesquisa pré-definidas.

O grupo de artigos obtidos como resultado foram subsequentemente sujeitos a avaliação de elegibilidade, como recomendado por Moher et al. (2009). Isto significa que os estudos que não cumpriam os critérios pré-definidos foram excluídos. Os critérios de inclusão foram: O trabalho académico tinha de ser uma referência primária de uma revista científica, escrita em inglês e a amostra analisada tinha de ter uma origem associada à comunicação nas social media online.

Posteriormente, os restantes 35 artigos foram transferidos para uma base de dados utilizando uma matriz de codificação. A matriz de codificação foi concebida para capturar os parâmetros-chave de cada estudo primário de uma forma padronizada e, portanto, comparável. Foi considerada informação geral, como o ano, localização e revista publicada, bem como informação temática específica, como o campo do turismo pesquisado e os meios analisados, juntamente com as categorias referentes à metodologia considerada, as ferramentas utilizadas e os resultados obtidos. A base de dados resultante foi então utilizada para obter declarações sobre a abordagem metodológica utilizada na análise da imagem de destinos online.

Como resumo dos resultados, foi desenvolvido um modelo conceptual, baseado nos conhecimentos obtidos a partir do grupo de artigos, que constituiu o conjunto de dados

para análise, para contribuir com um guião para a fase de preparação de uma futura investigação sobre imagem dos destinos online.

Em resumo, as principais conclusões são: TripAdvisor.com é a principal fonte para a análise da imagem de destinos online. Os investigadores recomendam a utilização da ajuda de software e linguagens de programação para a recolha e análise dos dados. À semelhança de estudos anteriores de *Destination Image*, os principais métodos aplicados na análise imagem dos destinos online são a análise quantitativa do conteúdo, a análise qualitativa do conteúdo e a análise dos sentimentos. Em combinação com a análise dos fatores cognitivos e afectivos, análise de co-ocorrência, e análise de correlação.

O presente estudo tem várias limitações. Que são: a perda de informação detalhada devido à redução dos estudos a parâmetros-chave comparáveis, a ausência de estudos anglo-americanos, devido à selecção do banco de dados, bem como a falta de testes de qualidade dos estudos incluídos.

Table of contents

Abstract.....	v
Resumo	vi
Lists of figures, tables, and abbreviations.....	ix
1 Introduction	1
1.1 Research Gap	2
1.2 Research Objectives	3
1.3 Research Relevance	3
1.4 Research Structure	3
2 Destination Image – From a traditional approach to a contemporary	4
2.1 Destination Image	4
2.2 Online destination image.....	6
2.3 Quantitative Content Analysis	8
2.3.1 Text mining.....	8
2.4 Qualitative Content Analysis	10
2.5 Sentiment Analysis	11
3 Methodology	12
4 Findings.....	19
4.1 Descriptive Analysis	19
4.2 Theoretical Analysis.....	22
4.3 Methods.....	25
4.4 Online DI studies knowledge gaps.....	29
5 Discussion and Future Research Agenda	31
5.1 Past and Present	31
5.2 Future research agenda.....	32
5.3 Conceptual Model and Practical guidance	34
6 Limitations	41
7 Conclusion.....	42
References.....	x
Appendix.....	xxvii

Lists of figures, tables, and abbreviations

Figures

Figure 1: FLOWCHART OF SLR PROCESS 13

Figure 2: PRISMA FLOWCHART OF ARTICLE SELECTION PROCESS 16

Figure 3: NUMBER OF PAPERS PUBLISHED PER YEAR 19

Figure 4: JOURNALS THAT PUBLISHED THE COLLECTED PAPERS..... 20

Figure 5: DESTINATIONS STUDIED DIVIDED BY CONTINENT..... 21

Figure 6: SENTIMENTS IMAGE OF CHINESE TOURISTS' REVIEWS GENERATED BY GEPHI 29

Figure 7: PHASES OF ONLINE DI RESEARCH DEVELOPMENT. CONCEPTUAL MODEL 36

Tables

Table 1: LITERATURE SELECTION CRITERIA..... 15

Table 2: SUPPORTING TOOLS FOR DATA COLLECTION AND ANALYSIS (PART 1 / 2) 26

Abbreviations

Abbreviation	Meaning
DI	Destination Image
DMO	Destination Marketing Organization
eWOM	electronic Word of Mouth
SLR	Systematic Literature Review
UGC	User Generated Content

1 Introduction

*“Die beste Bildung findet ein gescheiter Mensch auf Reisen /
Traveling is the best education for a clever person”*

(Goethe, 1796)

People have been traveling for centuries. Once for the purpose of trade and diplomacy and later increasingly also as a leisure activity. Not only Goethe but also many other poets and intellectuals like Mark Twain, Dalai Lama, and Fernando Pessoa describe the advantages of traveling. The travel market has been growing for years (Robino, 2019). Even distant destinations have become increasingly popular, at the latest since the introduction of low-cost airlines (Tkachuk, 2018). The United States of America, China, and Germany are the countries where most travelers originate from (Robino, 2019). Even though there is currently a sharp downturn in the tourism industry, as travel has been put on hold, due to the Corona pandemic, it is expected for tourism activity to recover again (Sharma et al., 2020)

The reasons for leisure traveling (push factors) can vary (Correia et al., 2007). Some people travel to find peace and relaxation, others want to experience adventures and gain new impressions, and some want to learn new things about culture and history (Correia et al., 2007). Accordingly, the destinations traveled to also differ to suit the respective needs of the travelers (pull-factors) (Kim et al., 2003). This explains why the list of the most popular destinations is so heterogeneous – some examples are London, Bangkok, Palma de Mallorca, Marrakech, and Bali (Robino, 2019).

But how do tourists decide which destination to visit and how do they know which one suits their desires? This is one of the key questions for marketers in the tourism industry and it has been intensively studied in tourism marketing research over the past 30 years (Baloglu & McCleary, 1999; Clarke & Hassaniien, 2020; Crompton, 1979; Gartner, 1986; Tasci et al., 2007).

One possible answer is given through the concept of Destination Image, which “can be defined as a person’s all sort of beliefs, thoughts and feelings (...) or a person’s overall perception about a destination and it is claimed that it considerably influences tourists’ travel decision” (Güneren et al., 2019, p. 1227). Sönmez and Sirakaya (2002) claim, that people are more likely to choose the destination they have positive thoughts and feeling

about. Understanding that, it is quite important for destination managers and tourist marketers to identify the predominant image of their destination and shape it in the desired way.

1.1 RESEARCH GAP

Nowadays the analysis of customers' perception of the Destination Image is mainly conducted online. In times of constant digitalization and the day-to-day use of social media platforms, companies have the opportunity to gather opinions and attitudes towards their products and brand online (Armstrong et al., 2014). These are easily accessible as users publicly share their opinions, recommendations, and experiences via various online platforms. A current market research showed that the world's most common social media platforms are Facebook, Youtube and WhatsApp (Chaffey, 2020). Additionally, platforms specialized in tourism allow users to search for information about destinations and allow tourism service providers to present themselves and their offers.

“For instance, Booking.com comprises 29 million accommodation listings in 154,000 destinations worldwide and operates in 190 countries. Similarly, TripAdvisor reports 8.4 million accommodation listings in 156,000 destinations spread across 49 markets, with 490 unique monthly visitors that provide more than 250 reviews and opinions per minute” (Lopez-Cordova, 2020, p. 7).

Several scientific studies have been carried out to observe the sentiments towards holiday destinations to support destination managers. For example Alonso-Almeida et al. (2019) carried out a content analysis on four major Chinese online travel agencies websites to measure the attitudes of Chinese tourists towards Barcelona. Micera and Crispino (2017) measured the sentiments toward Naples by analyzing a sample of online posts collected by the web-based social listening platform Blogmeter (<https://www.blogmeter.it/>), and Kim et al. (2017) tried to find out what negative feelings tourists have toward Paris by conducting a sentiment analysis on posts of the travel website *VirtualTourist* (an online travel guide site that does not exist anymore).

Nevertheless, at present there is no general overview drawn from previous studies on analysis approaches for measuring the Destination Image (DI) online through User Generated Content (UGC) (Cao et al., 2020).

1.2 RESEARCH OBJECTIVES

Therefore, the aim of the present study is to close this research gap and provide a systematic overview of the current situation in the research field of online DI analysis through UGC. The specific goals are:

To identify and explore extant state of the art of online DI analysis through UGC.

To examine and provide future research avenues addressing the DI online.

1.3 RESEARCH RELEVANCE

Following the proverb “learn from the past to prepare for the future” the relevance of the present research lies within the classification of gained knowledge into a larger context and thus facilitating orientation for future online DI studies. Since the research field of online DI, with an online dataset collection is still relatively new (Marine-Roig & Clavé, 2016), it can be helpful to have a unbiased systematic overview (Pickering & Byrne, 2014) of the methods used in the past to analyze DI online through UGC. As tourism marketing researchers encourage to further investigate in the field of online DI analysis (Mak, 2017; Marine-Roig & Ferrer-Rosell, 2018; Rahmani et al., 2019).

1.4 RESEARCH STRUCTURE

The following parts are structured in 6 sections, without considering the present introduction. Firstly, the theoretical foundations of DI will be presented. Secondly, the methodology applied in this study will be explained. Thirdly, the findings will be explored. Fourthly, the results will be discussed and put into context including an outlook for future studies. Fifthly the limitations of the research will be mentioned with suggestions for improvement and lastly a conclusion will be drawn.

2 Destination Image – From a traditional approach to a contemporary

To get a deeper understanding of the extant theoretical framework on DI and online DI, the following section explores its evolution and developments.

2.1 DESTINATION IMAGE

The concept of DI is one of the most widely studied topics in tourism management (Echtner & Ritchie, 1991; Hunt, 1975; Kim, 2014; San Martín & Rodríguez del Bosque, 2008). There are several different approaches to define and measure it (Madden et al., 2016). DI is based on the idea of brand image (Lopes, 2011). According to Keller (1993), brand image is the perceived attitude towards a brand and its attributes and benefits. A destination can also be seen as a brand that intends to convey a certain image. Gunn (1988) claims that the formation of the DI can be divided into three categories. As stated by the author, the first category is the *organic image* which describes the impressions tourists have of a particular destination without ever having visited it. This image can be based on organic sources without any commercial intent, like the Word of mouth (WOM) of family and friends as well as newspaper articles or television reports Gunn (1988). In the second category, the image is supplemented by marketing efforts of tourism authorities and travel agencies and is called *induced image*. The third dimension is called *complex image* and is the result of the actual visit to the destination and the experiences made there (Gunn, 1988). Other than in the first two stages, where the image is formed externally, the *complex image* is formed internally through the visit experience. Therefore, the image formed in the last phase is much more differentiated and realistic than in the previous phases (Echtner & Ritchie, 1991). Consequently, Marine-Roig (2019) defines the DI as the judgement an individual makes about a destination based on the evaluation of features, available activities and possible disruptive factors. These perceptions can be both positive and negative (Marine-Roig, 2019).

Echtner and Ritchie (1993) developed a model for DI that consists of three components. The *attribute-holistic* component, the *functional-psychological*, and *common-unique*. The intention was to show how the DI is composed of individual parts, which can be both tangible and abstract, to form a holistic image (Echtner & Ritchie, 1993).

One year later Gartner (1994) proposed that the DI results from an interaction of three related parts. He termed these components cognitive, affective, and conative. Affective and cognitive factors are known as one of the most important behavior shaping factors (Degen, 1999). Cognitive factors control the way in which behaviors are carried out. Cognitive stands for the reflective, the controlled, the reasoned, the rational (Degen, 1999). Typical examples are perception including judgment, information intake, learning, and memory (Roemer, 2014). In the context of DI this refers to the knowledge and beliefs a person has about a tourist destination (Pike & Ryan, 2004) and can be measured by the perception of different attributes of a destination. Examples of these attributes are *good shopping facilities, warm and friendly people, safety, varied and good food* (Calantone et al., 1989), *natural attractions, entertainment and events, historic attractions, travel infrastructure* (Chi & Qu, 2008), *quality of general tourist atmosphere, entertainment and recreation* (Assaker, 2014). The affective DI is an emotional response of individuals to a place (Basaran, 2016). The affective factors of human behavior can be divided into elementary and complex activating processes (Degen, 1999). Elementary processes include internal arousal and tension, and the basic needs of thirst, hunger, sleep, sex, and so on (Maslow, 1943). The complex activating processes describe emotions, motivations, attitudes, and values (Roemer, 2014). Emotion is directed inward toward the experience of the individual. Motivation is directed toward action and attitude toward an object (Degen, 1999). Values encompass many attitudes and thus relate to many objects and are more binding than attitudes. Values connect the individual with his/her social environment (Trommsdorff, 1989). Tourism and destination products are due to their nature emotion intense (Otto & Ritchie, 1996), consequently “tourist behavior researchers have been trying to understand in which shape and form and under which circumstances affective states influence tourists' future behaviors and their reactions toward brand evaluations.” (Duman et al., 2018, p. 4). To understand the affective attitudes towards a destination, destination management researchers often count and measure sentiments words (feelings) like *pleasant, relaxing, delicious, arousing, exciting, unpleasant, distressing*, etc. (Y. R. Li et al., 2015; Mak, 2017; Marine-Roig & Ferrer-Rosell, 2018). It is important to note that the cognitive DI significantly influences the affective DI (Lopes, 2011). In other words, “if tourists have positive cognitive, mental or intellectual evaluations and judgments about the destination’s attributes based on functional component, then they may have positive feelings or strong emotions about the destination” (Basaran, 2016, p. 173).

In Gartner`s (1994), three-step process knowledge, and belief about a destination (cognitive) are transformed into individual feelings and attitudes (affective), which then add up to an overall behavior towards the destination (conative). The final dimension in the triad, the conative DI, stands for the actions taken toward a destination, such as the likelihood of visiting the destination (Pike & Ryan, 2004), the revisit intention (Song et al., 2013) or a recommendation towards friends (Chen & Tsai, 2007). The conative image results from the two previous dimensions (Basaran, 2016). According to San Martín and Rodríguez del Bosque (2008) DI was traditionally analyzed through the cognitive approach focusing on tourists' images toward functional aspects of destinations (Duman et al., 2018) and only by time studies started to address the interconnections of cognitive and affective components in DI analysis (Duman et al., 2018). However by now Gartner`s theoretical construct from 1994 has been reproduced countless times in other studies (e.g. Assaker, 2014; Basaran, 2016; Chen & Tsai, 2007; Song et al., 2013) and offers a tool for empirically determining the DI.

2.2 ONLINE DESTINATION IMAGE

Since the shift of everyday life into digital space, and the increasing consumption of social media, more and more tourists share their experiences and opinions about holiday destinations openly on the internet (Schuckert et al., 2015). This electronic Word of Mouth (eWOM) - meaning noncommercial reviews and opinions - in combination with the marketing activities of the destinations, form a new information base for tourists (Lian & Yu, 2019). The online DI can thus be defined as "the online representation of the collective beliefs, knowledge, ideas, feelings and overall impressions of a destination" (Mak, 2017, p. 282).

The above mentioned theories on DI can be equally applied to online DI, with the exception that in this new information era, Tourist Generated Content about a destination is becoming one of the most important sources for users to gather information about the holiday region and contributes significantly to their travel planning (Lian & Yu, 2019). This content can be spread and found on different online platforms.

Either in blogposts from tourists on blogging sites (Marine-Roig & Clavé, 2016; Sun, 2020; Tseng et al., 2015), as comments or posts on advisor websites (Garay & Cànoves, 2015; Gkritzali et al., 2018; Oliveira et al., 2019), via interaction with the destination marketing organizations (DMOs) on social media like Facebook or Instagram (Gálvez-Rodríguez et al., 2020; Iglesias-Sánchez et al., 2020), or as postings of pictures using hashtags to refer to the destination (Deng & Li, 2018; Mak, 2017; Miah et al., 2017).

Depending on the platform, the contents often have a certain tonality. Moreover, since most platforms are for a global audience, comments in all different languages appear. This makes it easier to access information for people all over the world and give insights which languages groups (and thus tourists) are dominant.

Given that the first two phases of DI formation (cognitive and affective) (Gartner, 1994) are now influenced by other tourists – through online blogs, comments and photos – and not only by traditional media and DMOs (Mak, 2017), it is crucial for tourism marketers to capture and correctly interpret the prevailing online DI (Lin & Huang, 2006). In addition, it is useful for DMOs to know which attractions in their area are most values and in which activities tourists particularly to be engaged, in order to align themselves accordingly.

In the past, this information was often obtained in a costly and time-consuming manner through surveys and questionnaires. "By making use of big data, a DMO could obtain comprehensive insights on tourists' activities, their experiences and personal reflections" (Miah et al., 2017, p. 9). Therefore it is important to understand how online content can be evaluated in an efficient and structured way. Schuckert et al., (2015) found that the most common method for the analysis of online travel reviews is quantitative content analysis. This method then was extended by the inclusion of sentiment analysis (Cao et al., 2020; Chen et al., 2019; Kim et al., 2017, p. 20; Marine-Roig, 2019; Micera & Crispino, 2017). In the following both methods are introduced briefly.

2.3 QUANTITATIVE CONTENT ANALYSIS

Standardized quantitative content analysis is an empirical method that captures textual and visual media messages, as well as formal and content-related characteristics (Rössler, 2017)

The capturing is conducted by using predefined categories and their characteristics, which are written down in a codebook (Rössler, 2017). Developing a codebook or category system is the core task of content analysis. It is used to identify features from text units which are then assigned to the characteristics of the defined categories with the help of numerical values. Categories are comparable to the program questions of a survey. They are the exact definition of what is to be surveyed or measured. In effect, at the end of the category definition process, the characteristic values are determined (Brosius et al., 2009). The data can be analyzed descriptively in the further procedure. This reduces the complexity of the object of investigation to essential characteristics so that general statements can be made (Rössler, 2017).

Content analysis can be conducted manually or with the help of software. Unlike earlier studies with smaller, hand-selected data sets, scientists today work with enormous databases with thousands or millions of content pieces (Marine-Roig & Anton Clavé, 2016). Consequently, the focus is shifting more and more towards data selection and preparation. The actual content analysis is then performed by a programmed algorithm. The researchers Abburu and Babu (2013) therefore proposed a three-step-framework for web data extraction and analysis, which includes finding the URLs of webpages (web crawling), extracting the information from webpages (information extraction) and analyzing the data (text mining).

2.3.1 Text mining

Text mining, also called data mining, is a process for analyzing texts with the help of algorithms (Feldman & Sanger, 2007) A software using mathematical algorithms identifies the core information in a document. The method is often used to analyze large volumes of documents and publications that are related to a specific context. It can be used to prepare unstructured texts for database (Feldman & Sanger, 2007). Among other things, the software helps to select documents, identify patterns and trends, and to test hypotheses. Text files can also be automatically debugged, filtered, and corrected.

Semantic techniques enable the identification of key words for the respective context (Dickinger & Lalicic, 2016). In this way, relevant information can be extracted from large amounts of data with little time expenditure. Text mining involves working with unstructured data in pure text form. The content, logical structure and meaning are derived solely from linguistic and grammatical rules. Semantic and linguistic methods must be used to convert these into a form that can be read and evaluated by machines (Dickinger & Lalicic, 2016). The difficulty lies in translating these methods into algorithms, which enables programs to largely reproduce the intellectual activity of humans when reading texts (Feldman & Sanger, 2007).

The growing prominence of Big Data makes the application of text mining essential in many cases, since the larger the amount of data, the less human analysis is possible. An ever-increasing proportion of user-generated content on the web, also in the context of online DI, makes it necessary for DMOs and scientists to capture the meaning of postings by machine. Two supporting concepts often appear in the context of data mining which are co-occurrence analysis and clustering.

Co-occurrence analysis

“Co-occurrence analysis identifies ‘meaningful words’ based on the relative frequency of appearance with the target term” (e.g., the analyzed destination) (Költringer & Dickinger, 2015, p. 9). When two items co-occur, there is an association between the two entities. This is captured by comparing all sentences and documents related to the target term with content from the same website where the target term is not mentioned (Költringer & Dickinger, 2015). All paired items within a collection unit are then counted (Buzydlowski, 2015) and the significant words are identified by a chi-square test with Yates’ correction (Költringer & Dickinger, 2015). Through this technique topics can be identified that are related to the analyzed destination. For instance, shopping, beach, or the name of a famous museum.

Clustering

Clustering describes the grouping of data within a dataset. It is used in many scientific fields. In content analysis data it can be semantic clustering, which groups source artifacts that use similar vocabulary and divides them into different topics (Kuhn et al., 2007) or

geographical clustering to identify where the users are coming from or the popular areas of interest within a city (Miah et al., 2017). In the context of text mining clustering means programming a software to detect natural groups in data (Brownlee, 2020). Unlike classification – which is a supervised method to group information with the help of training data – clustering has no training data and is therefore an unsupervised method (Koch, 2020). In the context of online DI research clustering can help researchers to capture the most-mentioned objects and the connections between them and thereby identify the cognitive and affective perceptions in tourist reviews (Liu et al., 2019). These clusters could for instance show if the term beach appears in a cluster with words like sea, enjoying, surfing or with words like overcrowded, dirty and party.

2.4 QUALITATIVE CONTENT ANALYSIS

In addition to quantitative content analysis, qualitative content analysis is also carried out in DI research. This method originates from the scientific field of communication research. It was developed after the method of quantitative content analysis was criticized for being too superficial and not capturing latent content structures (Herkommer, 2012). The basic concept of Qualitative Content Analysis according to Mayring (2004) is to analyze texts systematically, and in that process to divide the material step by step, guided by theory, into categories that have been developed on the material. The aim is a summary, which represents the meaning contained in the text in so-called categories, which in themselves are organized in a system. The resulting system of categories, subcategories, and category definitions serves as a starting point for the interpretation of the text (Ramsenthaler, 2013). The systematic and theory-based procedure is designed to ensure quality criteria such as objectivity, reliability, and validity (Ramsenthaler, 2013). The main difference between qualitative and quantitative content analysis is primarily the evaluation. While quantitative content analysis focuses on statistical analysis by counting frequencies, qualitative content analysis refers in its evaluation to the context from which the categories were obtained. Thus, not only the manifest content can be captured, but also the latent meaning structures can be incorporated (Herkommer, 2012). Qualitative content analysis can also be supported by software. Examples for assisting software are Atlas ti, Leximancer or NVivo.

2.5 SENTIMENT ANALYSIS

Sentiments are groups of feelings or aggregated emotions (Liu et al., 2019). They are developed in a thoughtful socialized manner and not only out of pure instinctive emotions and tend to last longer (Munezero et al., 2014; Yu & Zhang, 2020). Identifying sentiments in vast amounts of unstructured content (e.g. user generated online content) is called sentiment analysis or opinion mining (Yu & Zhang, 2020). In travel and destination research sentiment analysis can be conducted to detect travelers' perceptions about holiday regions and thereby understand the strengths and weaknesses of this destination (Toral et al., 2018).

“Two approaches are widely used in sentiment analysis: machine learning and lexicon-based semantic orientation. Most studies used the machine-learning approach, owing to its high performance” (Ren & Hong, 2017, p. 2). The machine or better put computer program learns to classify with training data and then analyses the test data based on the previously built classifier (Ren & Hong, 2017). For this process a huge amount of content is needed to initially train the program (Alaei et al., 2019). “Machine learning algorithms reduce this high-dimensional feature space with the help of feature selection techniques which selects only important features by eliminating the noisy and irrelevant features” (Agarwal & Mittal, 2016, p. 1). In research fields which lack of sizeable training corpus scientist can make use of existing sentiment lexica to identify and categorize sentiments (Khoo & Johnkhan, 2018). These dictionaries can be created by humans, by machine or semiautomatically, meaning a combination of human and machine input (Alaei et al., 2019). The evaluation can also be done either mechanically as e.g. Gkritzali et al. (2018) do in their study on sentiments towards Athens, using the SentiStrength software (Thelwall et al., 2012) or manually like e.g. Hu et al. (2014) do in their study coding sentiments in text topics by hand.

Sentiment classification (Liu, 2015) meaning the grouping of text documents according to their sentiment classes (e.g. positive, negative, neutral) was conducted on three different levels in research of the past years (Ren & Hong, 2017). These are *content analysis* (Ariyasriwatana & Quiroga, 2016; Hu et al., 2014; Marine-Roig & Clavé, 2016), *document-based* sentiment analysis (Behdenna et al., 2018; Bucur, 2015), and *topic-based* sentiment analysis (Ficamos & Liu, 2016; Ren & Hong, 2017). As stated by Ren & Hong, (2017) all three techniques have advantages and disadvantages.

3 Methodology

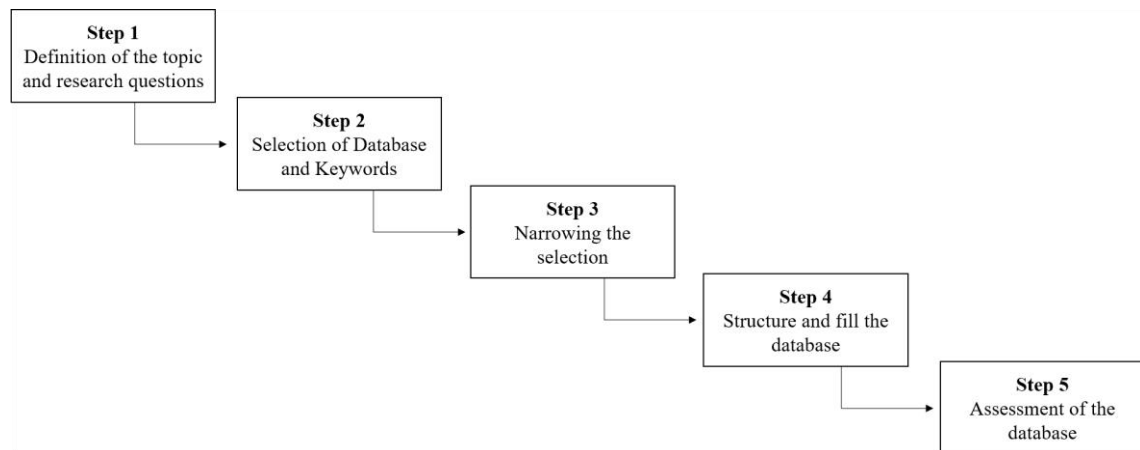
This study conducts a systematic domain-based structured review based on the recommendations of Pickering and Byrne (2014).

Systematic Literature Review (SLR) is “a scientific form of secondary research that relies on reproducible methods and procedures to identify, select, and appraise prior studies in response to a research agenda” (Lim et al., 2021, p. 536). SLRs are capable to provide a number of “critical discussions on a specific research theme by integrating extant literature, synthesizing prior studies, identifying knowledge gaps, and developing new theoretical frameworks” (Paul & Criado, 2020, p. 1). This particular style of literature review originates from the field of medicine (Bem, 1995; Tranfield et al., 2003) and was quickly adopted in all areas of science (Tranfield et al., 2003). Since then, it has been a frequently used method to obtain an overview of previously gathered knowledge and findings (Carter et al., 2015; Donnelly & Manolova, 2020; Nielsen et al., 2017). Authors like Fisch and Block (2018); Moher et al. (2009) and Pickering and Byrne (2014) provided recommendations and guidelines for conducting SLRs.

As stated by Tranfield et al. (2003) the advantages of SLR are that they are systematic, transparent, and reproducible in contrast to conventional literature reviews, which often lack thoroughness and rigor and are potentially biased by the researcher (Tranfield et al., 2003). In a systematic literature analysis, the criteria for the selection of literature are defined after a first fundamental review of the state of research. The selection of the literature is then based on the previously defined criteria.

Based on the recommendations of Pickering and Byrne (2014) a five-step process (which is presented in Figure 1) was applied to collect the relevant research contributions.

Figure 1: FLOWCHART OF SLR PROCESS



Source: own elaboration based on Pickering and Byrne (2014).

Step 1: Definition of the topic and research questions

First, a general review of the literature on online DI took place. The aim of this first literature review was to get an overview of the subject of DI and to elaborate a research topic and research questions. Furthermore, as suggested by Pickering and Byrne (2014), terms used in the research papers to describe DI were extracted in order to find the appropriate research keywords. From the finding in the general overview, described in the foundations section the defined research topic for this paper was:

“What are the existing theoretical frameworks on online DI research through UGC and which methods were used in the past in this tourism research field?”

In effect the questions aimed to be answered are:

- 1) When / were / for which purpose were the analyses of online DI conducted?
- 2) Which media / online platform were analyzed and in which language?
- 3) What was the size of the sample and how was it chosen?
- 4) Which methods / technologies to analyze the DI online were used and what are their advantages / disadvantages?
- 5) What were the outcomes of the studies and which recommendation for future DI studies were given?

Step 2: Selection of Database and Keywords

As a next step a data source from which the academic papers will be retrieved must be defined. To make the study reproducible it is useful to choose one or many academic databases for the data collection. There are plenty of academic databases for different scientific fields. Yang et al. (2017) stated that the seven major databases for tourism research are: Web of Science, ProQuest, EBSCO Host (Hospitality and Tourism Complete), Science Direct, Emerald, Scopus, and Sage. Moreover, Yang et al. (2017) found that prior tourism research reviews (e.g., Figueroa-Domecq, 2015; Mair et al., 2014) mainly used a combination of two or three out of these seven academic databases.

In the present paper the literature search was conducted utilizing three databases, namely: Web of Science, ProQuest and b-on (a Portuguese online knowledge library with over 16,750 scientific international publications from 16 publishers (b-on, 2020)), as they are the leading international database for tourism publications (Figueroa-Domecq, 2015), the largest, multidisciplinary, full-text database (Palmer, 2020) and a database with focus on the Iberian Peninsula.

To identify all relevant articles for the defined research topic in the databases and to ensure a reproducible literature review search keywords were defined. The aim of keywords is to “identify as much of the relevant literature as possible, but not extend too far into less-pertinent fields” (Pickering & Byrne, 2014, p. 543). The definition of the keywords resulted from the first general literature review on online DI.

All terms that are used in research papers to describe DI and its capture by mood analyses were identified, and thus the following search keywords were chosen: “DESTINATION IMAGE” linked with the Boolean operator “AND” to “SENTIMENT ANALYSIS” / “OPINION MINING” / “EWOM” as they often appear in keywords and abstracts of the relevant articles (Cao et al., 2020; Marine-Roig, 2019; Tseng et al., 2015).

Although the following SLR will focus on online DI it was chosen to name the keyword “DESTINATION IMAGE”, since the online DI can be seen as a amplification of the DI into the virtual world (Choi et al., 2007; González-Rodríguez et al., 2016; Marine-Roig, 2017). It was found that choosing “ONLINE DESTINATION IMAGE” as a keyword would narrow the search results too much from the beginning. The other three keywords were chosen since the process of detecting moods in online UGC towards a destination can be achieved by sentiment analysis (Ren & Hong, 2017), opinion mining (Chiu et al., 2015) or other forms

of the electronic word of mouth (eWOM) analysis (González-Rodríguez et al., 2016). The electronic databases were scoured for articles whose titles and / or abstracts contained the keyword “DESTINATION IMAGE” in combination with at least one of the other three keywords in the full text.

Step 3: Narrowing the selection

In August 2020, the search for the above listed keywords in all three databases produced 224 records. After exporting them to the research software Zotero (reference management software for bibliographic data and other research materials), the duplicates were sorted out and 146 articles remained. The remaining items were screened and checked using a specific selection criteria (Lim et al., 2021; Pickering & Byrne, 2014). Following the style of previous SLRs (Vada et al., 2020; Watson et al., 2018) a set of inclusion and exclusion criteria (Table 1) was determined to minimize potential bias.

Table 1: LITERATURE SELECTION CRITERIA

Inclusion criteria	Exclusion criteria
From scientific journal	Book, book chapter, conference proceedings, editorial, research note
Written in English	Written in other languages.
Primary source	Secondary source
Analyzed sample within the research must be from an open online source	Samples within the research only from questionnaires or personal interviews
	Articles older than 2010

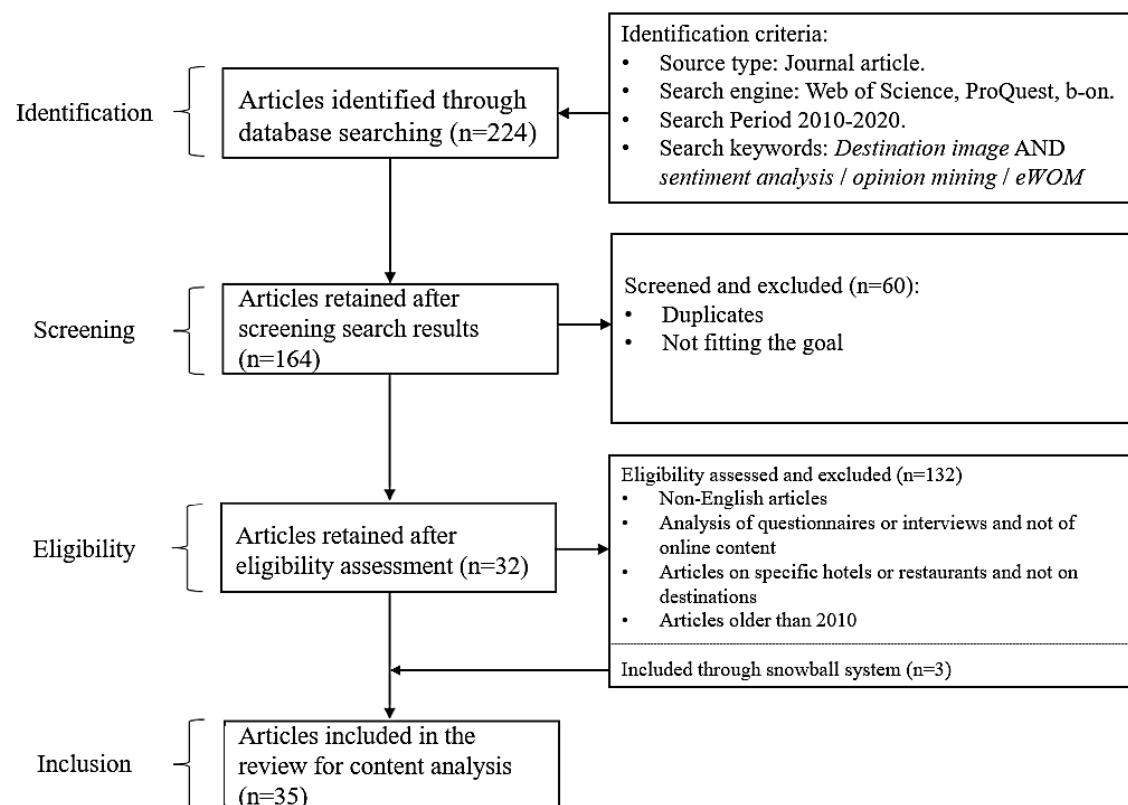
Source: own elaboration

To guarantee a selection of relevant and methodologically stringent literature, only articles from scientific journals were considered (Lim et al., 2021; Vada et al., 2020; Watson et al., 2018). The aim of this literature review is to get more insights on the methodology used in online DI research, with a clear focus on the methodology of using existing UGC. Therefore, only papers were included that collected their dataset from open online sources (e.g., Facebook, TripAdvisor, Travel Blogs). and studies which gathered their data only from questionnaires or personal interviews were excluded.

To be consistent with other SLRs (Belarmino & Koh, 2020; Vada et al., 2020; Watson et al., 2018), it was chosen to include only research papers written in English, since too much relevant information could be missed when translating. Aligned with Marine-Roig and Clavé (2016) it was decided to exclude research that focus on one specific event or aspect of the destination. Since the interest of the present study is the online DI as a whole, research was only included when it focusses on at least two destination aspects.

The period for this dataset was set at 10 years, from 2010 to 2020, which is shorter than other literature reviews on DI (e.g. Li et al. (2015) or Picazo et al. (2019), who have chosen 20 years), but since the objective is to provide the state of the art on the subject, the more recent the literature, the better. The founding of Airbnb (airbnb.com) in 2008 and Instagram (instagram.com) in 2010, followed by other social media on the internet, continuously influenced online DI and tourism marketing (Lim et al., 2021) as the amount of UGC continues to expand (accommodation ads on Airbnb expanded from 0,12 million in 2011 to 7 million in 2019 (airbnb.com, 2021)).

Figure 2: PRISMA FLOWCHART OF ARTICLE SELECTION PROCESS



Source: Adapted from Moher et al. (2009) and Lim et al. (2021)

For the transparent reporting of the inclusion and exclusion process of research in the database, the help of PRISMA Flowchart (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) was used. This four-phase flow diagram was developed in the medical sector and is nowadays widely used in SLR of all academic fields (Liberati et al., 2009). From all articles considered, 132 articles were excluded in the eligibility assessment (Moher et al., 2009) as they do not meet the above listed criteria (Fig. 2). Examples for subsequently excluded articles are:

Jalilvand and Heidari (2017) *Comparing face-to-face and electronic word-of-mouth in destination image formation*, as in this study the analyzed dataset was based on a survey and not on UGC.

Kirilenko and Stepchenkova (2017) *Sochi 2014 Olympics on Twitter: Perspectives of hosts and guests*, as the objective of this study is to find out the sentiment towards the event and not the destination.

Additional three papers were detected from cross-referencing and snowballing from the before identified articles as suggested by Pickering and Byrne (2014) and added to the 32 remaining articles. Which makes a total of 35 articles that are included in the review.

Step 4: Structure and fill the database

For analyzing the content of the papers and getting it structured a database was created, based on recommendations from Pickering and Byrne (2014). The coding matrix developed in November 2020 was designed to capture the key parameters of each primary study in a standardized and thus comparable manner. It contains general information like the year, location and journal published in, as well as thematically specific information as the researched tourism field and the analyzed media, together with categories on methodology and results.

At first, six papers were entered in the database as a testing of the categories (Pickering & Byrne, 2014). The testing showed that an addition of the category “THEORETICAL OR PRACTICAL ADVICE” is useful for the further analysis. A great challenge was to record the central findings as uniformly as possible, since the range of measurement instruments and statistical methods used in the original studies was extremely heterogeneous. Subsequently all 35 articles were entered into the database.

Step 5: Assessment of the database

As advised by Pickering and Byrne (2014) the last step is a careful evaluation of the filled database, including a counting and grouping of the results, to understand which methods, findings, etc. are the most important. This helps to describe the overall picture resulting from the research. At the same time the “results need to be related back to the overall topic and the original research questions” (Pickering & Byrne, 2014). In order to obtain an overall picture of the results in the database, the entries first were examined and filtered under various aspects. This made it possible to draw more general conclusions about the temporal and local distribution of the studies and research fields.

The database then was subdivided into smaller summary tables, to gain deeper insights into the methods used, the supporting software, and the outcomes and findings. These smaller summary tables were first described and analyzed individually before being correlated in the final step to obtain an aggregated picture. This helps to draw conclusions about the methods and results and supports a more comprehensive analysis of the topic. The assembled database is attached in the Appendix 1.

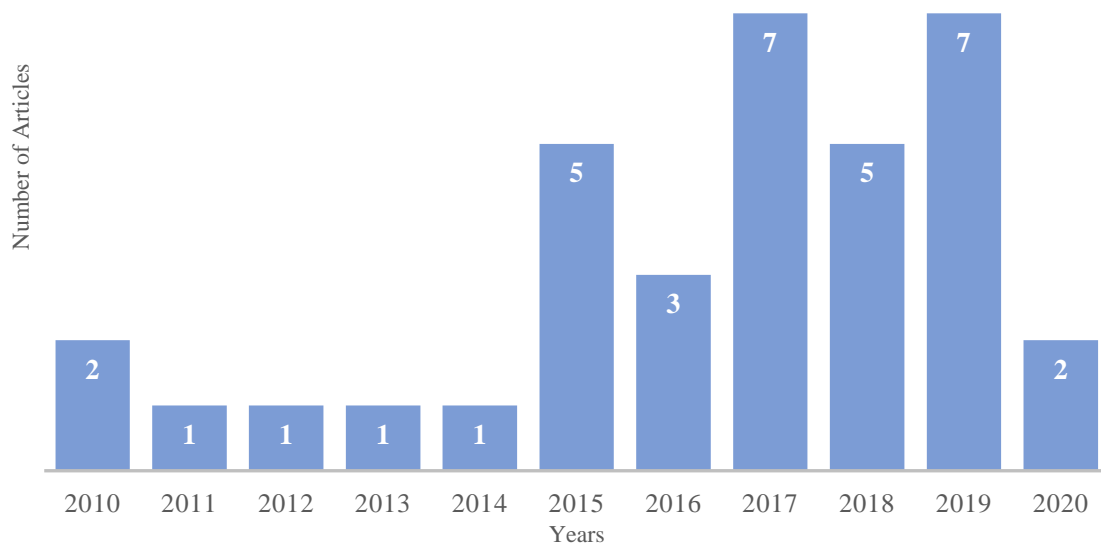
4 Findings

After conducting a SLR on research performed in online DI analysis, utilizing the three academic databases Web of Science, ProQuest and b-on, as well as a collection of papers through the snowball system, 35 empirical research articles from 22 different journals published between 2010 and 2020 were identified as relevant.

4.1 DESCRIPTIVE ANALYSIS

The results of the database consisting of 35 papers contains at least one article per year (Fig.3). Most articles on Online DI were published in the years 2019 and 2017 (seven each) followed by 2018 and 2015 (five each). Which makes a total of 77% of the articles in the years 2015-2020 and indicates the topicality of the subject and the still young field of research.

Figure 3: NUMBER OF PAPERS PUBLISHED PER YEAR

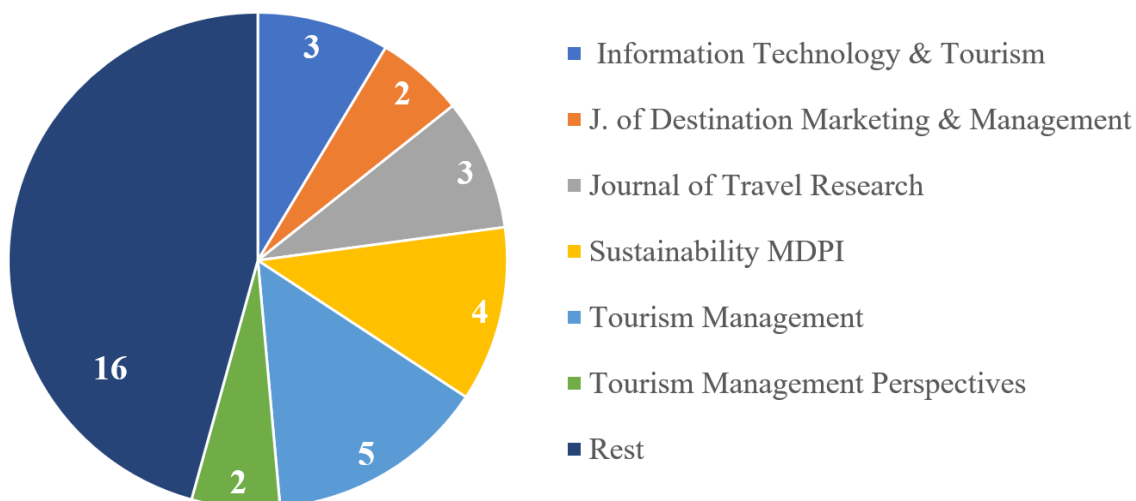


Source: Own elaboration

The database includes research papers retrieved from 22 different scientific journals. As shown in Figure 4, *Tourism Management* published five (23%) of the analyzed articles and *Sustainability* four (18%), which makes them the main journals on this topic.

While most (73%) journals are related to Tourism, Travel, and Hospitality, six journals in the data set are from other scientific fields such as Technology, Social Science, Communication Science and Business Management.

Figure 4: JOURNALS THAT PUBLISHED THE COLLECTED PAPERS



Source: Own elaboration

The researchers who published most articles in the given period from 2010 till 2020 are Estela Marine-Roig, with five publications (14%) in the years 2016-2019, followed by Luís Garay, who published three articles (9%) in 2015, 2017 and 2019, and Yi Lui with three publications in 2019.

Regions and Languages

The research papers included in the database explored the online DI in Africa, America, Asia, Australia, and Europe (Fig. 5), with a clear dominance of destinations in Europe with 17 articles (49%) and an absence of destinations in South America.

The countries which DI researchers explored mostly are Spain and China, with seven papers (20%) for each country. Studies in Europe mainly focus on cities such as Barcelona, Vienna, and Paris, whereas studies from other continents take a broader look on whole regions (e.g., East Taiwan or Zanzibar Island). As mentioned in the methodology part the included studies had to analyze at least two tourism aspects. The analyzed tourism aspects within the destinations are quite heterogenous. While 29% of the studies took a closer look on city and culture among other aspects, 22% considered restaurants in their DI. From the 35 papers in the database 19% took hotels into account when analyzing the DI and 16% the aspect of beach and relaxation tourism. Besides sports (considered by 13% of the studies) other additionally examined topics were shopping, nature, infrastructure, and transportation, such as safety and people.

Six of the 35 research papers explicitly focus on the online DI perceived by Chinese tourists. For instance *Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations* (Liu et al., 2019) or *Online representation of Switzerland as a tourism destination: An exploratory research on a Chinese microblogging platform* (Hu et al., 2014).

In the distribution of languages from UGC content analyzed Chinese is also strongly represented. 44% of the studies considered online content in English and 28% explored Chinese comments and blogs. Only one study each (3%) focused exclusively on Spanish and Italian comments.

Figure 5: DESTINATIONS STUDIED DIVIDED BY CONTINENT



Source: Own elaboration

Analyzed Media

Regarding the media source the most prominent were found to be Advisor Websites, which are used as the sample content in 26 studies (74%). From these 26, 14 academic papers analyzed content retrieved from TripAdvisor.com which makes this website the most popular content source for DI studies. 19 studies (54%) used Travelblogs to gather content, especially from the blog-websites TravelBlog.org and TravelPod.com. Only four studies (11%) collected their content from Twitter and even less from Facebook (9%), Instagram and YouTube (two (6%) each and only in combination with other online sources from social listening platforms).

The approach of media source choice differs from study to study. The main reason for researchers to choose TripAdvisor is the size and the worldwide leading status of the website (Gkritzali et al., 2018; Marine-Roig, 2019; Oliveira et al., 2019; Van der Zee & Bertocchi, 2018). In her study 2017 Marine-Roig, even applied a “weighted formula of aggregations of rankings”(Marine-Roig, 2017, p. 7), to verify that TripAdvisor is the leading tourism website in terms of visibility, popularity, and size. Another popular website for analysis was Ctrip.com, a Chinese tourist review website and provider of travel services. This website was mainly chosen when analyzing Chinese Tourists, since it is one of the biggest tourism websites in China (Ren & Hong, 2017; Sun, 2020). Researchers who have chosen travel blogs as a source point to the holistic perspective that this medium brings. Users can post texts as well as pictures and videos on the topics of their choice (Mak, 2017; Rahmani et al., 2019; Tseng et al., 2015).

The ratio of studies that selected only one media source to studies with multiple media sources is very balanced (17 to 18). Govers (2012) and Micera and Crispino (2017) decided to include all types of social and online media. With the help of so-called social listening tools (in this cases Radian6 and Blogmeter) the researchers crawled the entire internet for their defined keywords. The collected data is then grouped and analyzed by the tool. Researchers who have analyzed only one social media platform claim this approach as one of their major limitations, especially when the analyzed media is not TripAdvisor (Bultjens et al., 2013; Garay, 2019; Garay & Morales Pérez, 2017). But even some researchers that have used TripAdvisor as the only source of their research, claim that a comparison between more than one platform would be useful (Marine-Roig, 2019; Van der Zee & Bertocchi, 2018).

4.2 THEORETICAL ANALYSIS

Among the theory-driven findings, the following commonalities can be identified: UGC clearly differs from other sources – e.g. surveys or media reports – in determining DI (Bultjens et al., 2013; Dickinger & Lalicic, 2016; Garay & Cànoves, 2015; Költringer & Dickinger, 2015; Marine-Roig & Ferrer-Rosell, 2018). According to Költringer’s and Dickinger’s (2015) findings, UGC are the more diverse and richer source in identifying DI. Especially in distinguishing emotion-based experiences, statements from UGC differs significantly from those from surveys (Dickinger & Lalicic, 2016).

In addition, not only does the UGC contrasts to other media sources but also there are variations within the UGC which should be kept in mind. The results of Liu et al. (2019) show that Chinese tourists' perceptions and preferences differ from those from other international tourists, and Jani and Hwang (2011) found that the DI of tourists who have already visited the destination is different from those who have not yet been there. Marine-Roig and Ferrer-Rosell (2018) are concluding from their findings that perceived DI differs strongly from projected DI. Therefore, Tseng et al. (2015) advise to separate the DI from the writers of posts from the DI of the readers.

Moreover, in their findings the researchers provide several frameworks and guidelines for future studies on DI. Marine-Roig and Anton Clavé (2016) propose a detailed model for content analysis of large amount of data, with significant insights concerning both the cognitive and the affective components of DI. In her study three years later Marine-Roig (2019) present an initial integrated framework for analytics on tourists DIs from a massive amount of UGC. Another framework was presented by Ren and Hong (2017) for the efficient use of Latent Dirichlet Allocation model (a statistical approach for topic modeling). The framework includes word-topic and document-topic matrix.

Miah et al. (2017) developed an IT artefact in the form of a general method for generating meaningful information and predictive insights from geotagged photos. In the same year Marine-Roig (2017) presented and applied a method design to measure the perceived and transmitted DI from online travel reviews presented in search engines. Focusing on negative sentiments towards a destination Kim et al. (2017) developed a hybrid method model, including the analysis of travelers' online review data by employing both sentiment analysis methods for identification of potential problems and co-occurrence analysis for solution suggestion. As stated before, in many cases researchers found a gap between perceived and projected DI. Therefore, Marine-Roig and Ferrer-Rosell (2018) established a methodology to asses these tourists DI gaps in relation with the DI components.

Categorization

Regarding the categories to analyze DI many study used established destination attributes, such as Sights, Activity, Accommodation, Environment, Infrastructure, just to name a few (Jani & Hwang, 2011; Költringer & Dickinger, 2015). At the same time some researchers developed new categories suitable for their destinations analyzed. These are *travel phenomena & stories, willingness to go, effective official communication, link sharing* in the case of Hu et al. (2014) analysis of the Chinese DI of Switzerland. Marine-Roig and Clavé (2016) developed eight categories of attraction factors for the region of Catalonia, namely: *food and wine, intangible heritage, leisure and recreation, nature and active tourism, sports, sun, sea and sand, tangible heritage, and urban environment*. Just like Gkritzali et al. (2018), who created five DI categories to measure the DI of Athens after the crisis, which are: *safe environment, hospitable environment, general mood and atmosphere, relaxing effect, and authenticity of experience*.

Timeframe and sample size

There are significant differences in sizes and time span of the collected samples. Marine-Roig and Anton Clavé (2016) chose the longest time span for their data collection with eleven years, followed by Deng and Li (2018) and Gkritzali et al. (2018) with ten years each. Other researchers such as Micera and Crispino (2017) and Law and Cheung (2010) considered a much shorter time span of only four months for the collection of data. Depending on the time span chosen, a current snapshot of the DI is taken or a development over the years, including events and political changes is observed.

The amount of TripAdvisor reviews analyzed varies from 472 reviews (Oliveira et al., 2019) up to over 300,000 reviews (Marine-Roig, 2019; Marine-Roig & Ferrer-Rosell, 2018; Van der Zee & Bertocchi, 2018). The span of the number of analyzed photos is also quite large and rises over time; from just under 2,000 photos in the smallest sample (together with textual analysis) (Mak, 2017) over about 21,000 photos on Flickr (Deng & Li, 2018) up to a little over 238,000 photos (pure photo metadata analysis) (Miah et al., 2017).

Again, the researchers used different approaches to generate their samples. While some downloaded all available posts on the selected destination in the specified time period (Marine-Roig, 2019), or used a web crawler with the specification of a certain sight as

search criterion (e.g. Van der Zee and Bertocchi, (2018): search criterion Antwerp Train station), others additionally considered certain aspects in their pre-selection, such as seasonality (Garay, 2019) or the rating on TripAdvisor (Dickinger & Lalicic, 2016). From the researchers who chose Travel Blogs as a media source, some made the selection of individual posts with the help of search engines. For instance Hu et al. (2014) used the selected destination as search keyword in a search engine (in this case Sina Weibo) and then included all posts from Travel Blog pages in the fitting timeframe in their dataset.

4.3 METHODS

The main methods applied in the analyzed studies were quantitative content analysis, qualitative content analysis, and sentiment analysis. This equals the approaches of earlier DI studies. However, 29 of the analyzed studies (83%) used a combination of at least two methods, to gain insights on the online DI. While the quantitative analysis mainly went along with sentiment analysis and the examination of cognitive and affective factors, sentiment analysis as a main method was mainly accompanied with co-occurrence analysis.

Due to the vast amount of online data available, it is becoming more common among researchers to use the help of software and programming languages for data collection and analysis. As such 27 (77%) of the 35 studies in the dataset claim to have used the help of these tools, which are listed in Table 2.

Table 2: SUPPORTING TOOLS FOR DATA COLLECTION AND ANALYSIS (PART 1)

Technology / Programme	Purpose	Researchers	Year
Atlas ti	Qualitative content analysis	Sun et al.	2015
Bazhuayu	Website crawler for data collection	Ren & Hong	2017
Blogmeter	Web-based social listening tool for data collection and clustering over all social media channels	Micera & Crispino	2017
CAQDAS	Computer-assisted qualitative data analysis software	Sun et al.	2015
CATA Software	Computer-aided text analysis	Garay & Cànoves	2015
CATPAC II software	Textual content analysis	Mak	2017
GATE (General Architecture for Text Engineering)	Text processing	Miah et al	2017
Gephi	Clustering and co-occurrence analysis	Alonso-Almeida et al. Liu et al.	2019 2019
Kimono (web-based scraping software)	Finding reviews on Antwerp and match them to user profiles	Van der Zee & Bertocchi	2018
Language Detection program	Detecting the language of the blog entries	Marine-Roig & Clavé	2016
Leximancer software	Text mining	Oliveira et al. Tseng et al.	2019 2015
LIWC	Linguistic inquiry and word count, to separate reviews into a set of words	Alonso-Almeida et al. Liu et al.	2019 2019
Ncapture®	Nvivo add-on tool for capturing and structuring tweets (Twitter)	Garay	2019
Nvivo Software	Visual content analysis and mise en scene analysis	Mak Sun	2017 2020
Python	Data collection / word segmentation and part-of-speech tagging / visualization of the word frequency / to calculate the score of positive and negative tourism emotions / co-occurrence analysis	Chen et al. Kim et al.	2019 2017
Radian6	Social listening tool for data collection and clustering over all social media channels	Govers	2012

Table 2: SUPPORTING TOOLS FOR DATA COLLECTION AND ANALYSIS (PART 2)

Technology / Programme	Purpose	Researchers	Year
SAS (Enterprise Miner Workstation)	Text topic modeling	Rahmani et al.	2019
SentiStrength software	Sentiment analysis	Gkritzali et al.	2018
SentiWordNet	Measuring the sentiment values of adjectives	Deng & Li	2018
Site Content Analyzer software	Conducting keyword count	Marine-Roig & Clavé	2017
Stanford sentiment analysis tool	Sentiment analysis	Kim et al.	2017
SURF (Speeded-Up Robust Features)	Feature descriptor for photo content	Miah et al	2017
TDIV programme	Tourist DI visualization	Cao et al.	2020
Text STAT software	Calculating word frequencies	Jani & Hwang	2011
ThoughtView software	3D perceptual maps for the TGC and the NTO datasets	Mak	2017
UCINET	Conducting the social network analysis and visualization	Van der Zee & Bertocchi	2018
UltraEdit	Extracting meta data	Marine-Roig	2017
webLyzard crawling agent	Data collection	Költringer & Dickinger	2015

Source: own elaboration

The tool Python was used in four studies (11%) for the analysis of TripAdvisor, Flickr, virtualtourist.com (an online travel guide site that does not exist anymore) and several Chinese tourism websites. Python is a programming language developed in the 90s, which enjoys great popularity due to its ease of learning. Programming languages are formal languages with which computing rules can be defined. These rules then can be executed by a computer. Python can be extended by special plug-ins and add-ons and thus obtains a multitude of application possibilities (Augsten, 2019). In the reviewed literature it was found to be used to crawl the reviews and comments, for co-occurrence analysis, for visualization of word frequencies in word clouds and to calculate the score of positive and negative tourist emotions (Chen et al., 2019; Deng & Li, 2018; Kim et al., 2017; Sun, 2020).

Oliveira et al. (2019) and Tseng et al. (2015) use the help of Leximancer software for mining of their data. Leximancer is a software, which identifies concepts in textual data and then visualize them in concept maps, to have the key ideas at a glance. Unlike other content analysis software Leximancer is working with its own algorithms for identifying the main ideas within passages, instead of applying term coding or word frequency count (Tseng et al., 2015).

Garay (2019) and Garay and Morales Pérez (2019), who analyzed the DI of Spain via Twitter, uses the NVivo software, more precisely NCapture, to mine Twitter posts. NVivo is a software which helps researchers to find and organize unstructured non-numerical data. The web-browser extension NCapture helps to collect web content and import it to NVivo. It captures the data as PDF or as data sets, which enables the researcher not only textual data but also photo and video content from social media (*NCapture*, n.d.). NVivo software was also used by Mak (2017) for a visual content analysis and a mise en scene analysis on travelblogs.

Another tool is Gephi, which was used for data collection and analysis by Alonso-Almeida et al. (2019) and Liu et al. (2019). Gephi is a network analysis and visualization software, that is applied in science and journalism (*Gephi - The Open Graph Viz Platform*, n.d.). In the mentioned researches the software has helped to collect the most frequent words in reviews and to build a cluster network through co-occurrence analysis (Alonso-Almeida et al., 2019; Liu et al., 2019). The results are then visualized by Gephi as a clustered network as presented in Figure 6. Similar to the visualization in Leximancer,

components, since their goal was to get a deeper understanding of the affective motives. Nevertheless, they are recommending future researchers not to follow their path but to consider all DI components in the research process, as it provides more holistic insights.

Marine-Roig and Ferrer-Rosell (2018) recommend to future researchers to incorporate the aspect of familiarity in their research by distinguishing between first-time visitors and repeat visitors.

Hu et al. (2014) see their limitations in the amount of analyzed sources and a relatively small range of keywords used. As their dataset only considered microblog textposts from Sina Weibo and they suppose that other keywords might have provided other relevant data. For future researchers they recommend applying other additional methods, e.g., questionnaires, for further interpretation of the results. Similar to this Tseng et al. (2015) recommend a mixed method approach for future researchers. Moreover, they see a need in clarifying the differences in between the DI and the Country Image of Mainland China as well as in a comparison of the DIs from short-haul and long-haul international tourists and different cultural groups.

A knowledge gap in online DI research was found by Rahmani et al. (2019) concerning common memory and the formation of an objective DI. The researchers state that therefore a large group of tourism researchers would need to explore large-scale qualitative data.

Moreover, Garay and Cànoves (2015) show the need to further analyze the characteristics of the different social media platforms to better interpretate the findings. This includes the actors, the behavior patterns, and the components of the specific platforms, as they can vary significantly.

5 Discussion and Future Research Agenda

The main objective of this study was to extract and cluster findings for future research on DI. Therefore 35 research papers on online DI published between 2010 and 2020 were analyzed.

5.1 PAST AND PRESENT

It was found that the journals which published the most articles on this topic are *Tourism Management* and *Sustainability MDPI*. The main authors in the field of online DI analysis through UGC in the years between 2010 and 2020 are Estrela Marine-Roig, Luís Garay and Yi Liu. Europe and Asia were the two continents where most researchers lied focus on. Within Europe the DI online of Spain was the most analyzed and within Asia the DI of China. The small number of research contributions on America is most likely due to the databases used in this study, which are Web of Science, ProQuest and b-on. Looking at the tourists, it becomes apparent that in the last decade a main focus in research was on Chinese travelers. Which is understandable, having in mind that China is the nation sending out the second largest number of tourists in the world.

Within Europe mainly Metropoles like Barcelona, Paris or Vienna where analyzed, in conclusion the main tourism aspect studied in the last decade are City and Culture.

To gather the UGC for online DI analysis advisor websites and travel blogs are the most common sources. The rating platform TripAdvisor is by far the most popular website among tourism scientists to collect data for an online DI analysis. Concerning the time span and size of the analyzed sample no clear trends could be identified. The time span for sample collection varies from four months to ten years and the sample size varies from less than 500 up to over 300,000 UGC items. Probable reasons for this large spread could be the number of available comments or reviews and the way of collection. Since some destinations are less more popular than others and therefore the amount of available content online is smaller. As well as researchers collecting the data manually have less capacities that researchers collecting the data with the help of software. Therefore, the usage of software as a support for online DI becomes increasingly popular. The SLR analysis showed that since 2017 most researchers are incorporating software support in their research, for instance for data collection, word count, data clustering and sentiment analysis. The main methods applied for analyzing online DI through UGC are quantitative

content analysis, qualitative content analysis and sentiment analysis. In combination with the examination of cognitive and affective factors, co-occurrence analysis, and correlation analysis. This is similar to the methodological approach of former studies of DI.

Among the analyzed studies different emphases can be seen in the descriptions of the results. While some studies present their findings in a very practical and application-oriented way, applicable for destination managers of the region, other studies focus on theory-driven findings, which include more general insights for tourism research. For example, Garay (2019) provides tips for future DI studies by recommending creating stakeholder profiles, breaking down attributes into cognitive and affective, and considering seasonality when determining DI. On the other hand, Alonso-Almeida et al. (2019) specifically focus on Barcelona when describing their findings, in which they explain the behavior of Chinese tourists in the city as well as the positive and negative aspects about Barcelona from a Chinese perspective.

5.2 FUTURE RESEARCH AGENDA

In the following recommendations that were given in the studies are presented in clusters.

Micera and Crispino (2017) state that it is crucial to prior develop a strategic reflection on goals, objectives, and expected results. Due to the immense amounts of data available and the many possibilities of analysis tools, a clear roadmap is indispensable in order not to get lost in the data volume. The following studies present models that can serve as a guide for the general approach and the development of the strategy:

Marine-Roig and Clavé, (2016) *A detailed method for destination image analysis using user-generated content*. The researchers describe in a detailed way each step of online DI analysis with a HTML dataset by conducting an analysis of Catalonia with large amount of UGC.

Kim et al. (2017) *What makes tourists feel negatively about tourism destinations? Application of hybrid text mining methodology to smart destination management*. Short overview of the research process of UGC analysis on destination hospitality services. Conduction of a post-hoc co-occurrence analysis, for a deeper understanding of the reasons of negative feelings of tourists.

Marine-Roig (2019) *Destination Image Analytics Through Traveller-Generated Content*. Holistic view on DI analysis provided with a model to measure tourist DIs including aspects of sustainability, satisfaction, and loyalty of visitors.

For sentiment analysis, the authors of several studies advise to examine the texts to be analyzed not as a whole, but to divide it into topics and then conduct a sentiment analysis for the individual topics (Hu et al., 2014; Rahmani et al., 2019; Ren & Hong, 2017; Sun, 2020). Since reviews or blog posts usually cover a wide range of subjects and sentiments, no clear picture can be identified in the holistic view, or rather the positive and negative impressions blur into a neutral overall picture. However, if these texts are divided into topic groups, it is much clearer to see how tourists feel about the different aspects of the destination. In addition, the individual topics can be compared with each other and ranked by positive and negative. In its conclusion, an overall picture of the destination can be compiled from this.

As indicated before, there are significant differences between UGC and DMO content. The differences lay mainly in the expression of feelings and emotions, which are considerably more represented in UGC than in other sources. Chen et al. (2019) state that opinions in UGC are more positively biased – because of the so-called Pollyanna effect, which is the tendency of people to recall positive and pleasant information easier than negative one. Termed by the US psychologist Charles Osgood in 1964 (Doyle, 2011) – but this is not confirmed by other studies. However, the found differences by media format indicate a gap between projected and perceived DI (Marine-Roig & Ferrer-Rosell, 2018; Tseng et al., 2015). It is therefore advisable to take a close look at the deviations in order to narrow this gap by adjusting tourism offers and marketing strategies. The smaller the gap between expectation and reality, the greater the tourist's satisfaction (Oliver & Winer, 1987; Weber, 1997). The emotions expressed in UGC can be specifically incorporated into DMO materials to create a more vivid image of the destination (Költringer & Dickinger, 2015). Mak (2017) and Deng and Li (2018) for instance, show approaches to incorporate tourists' photos into DMO material to visually depict emotions linked to the destination. Additionally, destination managers could repost tourists' online content strategically on their social media channels, to create a closer bond with the tourists but also to show *real live* pictures of the destination.

The findings of this study reinforce the need to further investigate online DI using UGC (Mak, 2017; Marine-Roig & Ferrer-Rosell, 2018; Rahmani et al., 2019). For instance, Mak (2017) promotes to explore the online DI through photographic UGC. Especially the differences between different social media photo content platforms are poorly investigated so far. Marine-Roig and Ferrer-Rosell, (2018) propose to compare the DI of similar destinations to filter what makes each destination really unique. Rahmani et al. (2019) generally urge tourism researchers to do more studies on large scale data. Since the comparison of several large-scale data studies can than answer the question of “how objective destination images can form and whether common memories (can) exist” (Rahmani et al., 2019, p. 12)

5.3 CONCEPTUAL MODEL AND PRACTICAL GUIDANCE

To summarize the findings of this study a conceptual model as presented in Figure 7 was developed, based on the insights of the studies in the dataset. It can be seen as a guidance for the preparation phase of future DI research.

The conceptual model is divided into four phases – selection , collection, analysis and reporting – which was adapted from other researchers Kim et al. (2017) and Marine-Roig (2019). The division indicates that already in the theoretical preparation of an online DI research decisions for all for phases should be made.

Before diving into the individual phases, researchers should define the strategic goals for their research. This can be supported by answering the following questions:

- What is the purpose of the research?
- Does the research has a theoretical or practical focus?
- Which target group is aimed to be analyzed?
- Which destination aspects are seen as important?

Selection Phase

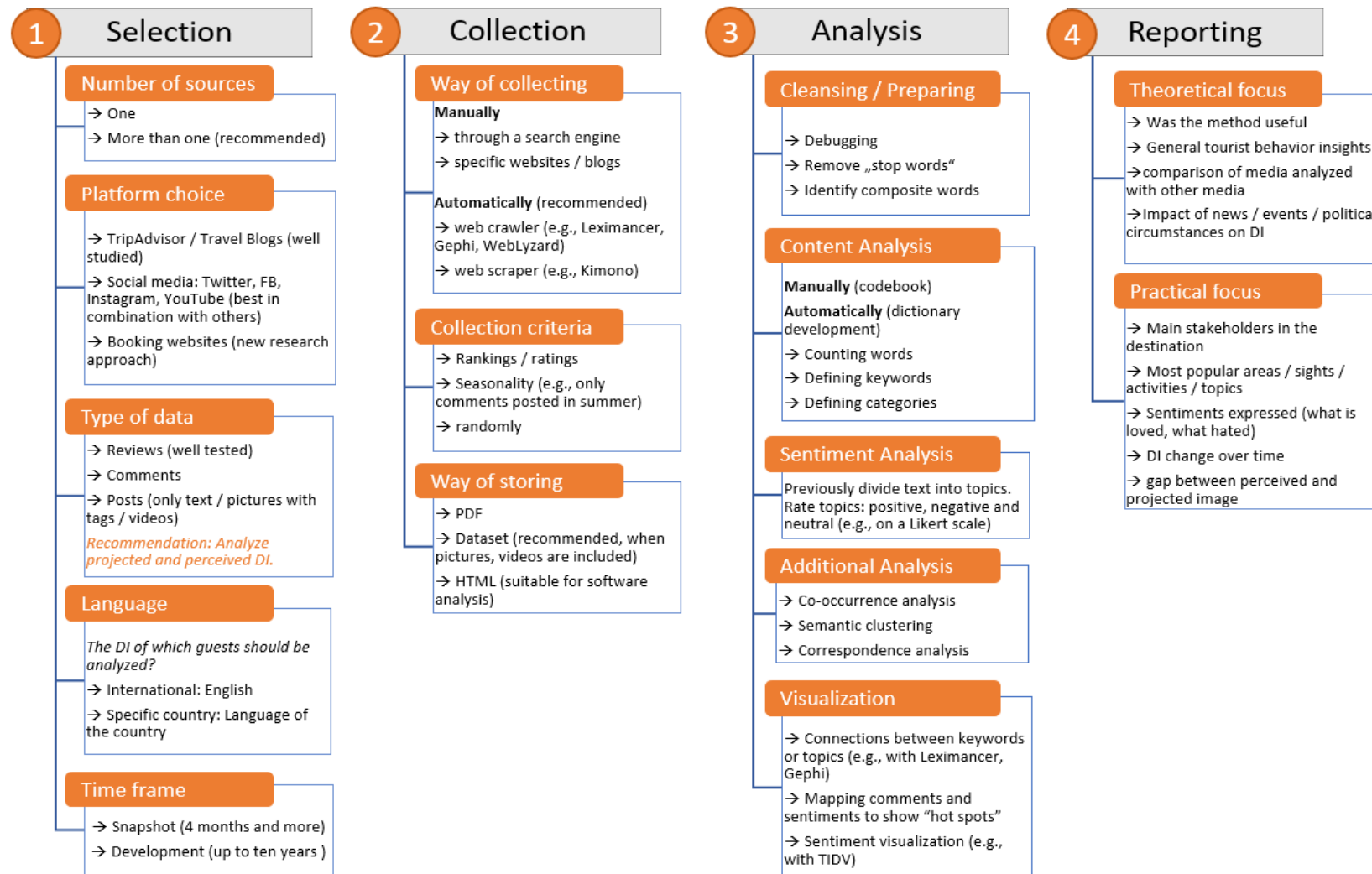
In the selection phase it should be defined, not only which destination is going to be analyzed but also from which perspective. According to the choice of perspective the fitting media can be selected. It is crucial to understand, that the DI will vary depending on the analyzed source. To get a more holistic DI it is therefore recommended to choose more than one media source.

In the dataset studies, the main media / online platforms analyzed are TripAdvisor, Travel advisor websites and Travel Blogs. None of the studies included a booking website, such as Booking.com, Hostelworld.com or Airbnb in their media sample. Even though those websites combine the projected and perceived DI since they include general information about the destination together with comments and reviews of guests. This could be a potential research field of upcoming studies.

Especially the comparison of two accommodation booking websites, with different target groups (e.g., young travelers and luxury travelers) could give a holistic DI from various tourist perspectives. In the analyzed studies researchers mention that the users writing travelblogs and reviews online are limited to a younger age group (Sun, 2020). However, since the internet is becoming more and more accessible to every age group and the users become more diverse, this aspect should be kept in mind but not be seen as a limitation.

The question of platform choice should relate to the question of strategic goals. Meaning, if the strategic goal is to deliver substantial DI insights for destination managers opt for classical platforms. If it is to present new insights in the field of tourism research, choose unconventional platforms.

Figure 7: PHASES OF ONLINE DI RESEARCH DEVELOPMENT. CONCEPTUAL MODEL



Source: Own Elaboration

After the choice of media platform, the data type must be selected. For instance, reviews, comments to certain sights or topics, posts in discussion forums or travel reports. Moreover, it must be decided whether only text data is considered or if pictures and videos are also going to be included. In past research, a significant gap between projected and perceived DI was found (Buultjens et al., 2013; Dickinger & Lalicic, 2016; Garay & Cànoves, 2015; Költringer & Dickinger, 2015; Marine-Roig & Ferrer-Rosell, 2018). Hence, it is recommended to consider content from both sites, projected (DMO content and media) and perceived (UGC from tourist, who have already visited the destination).

The language of the collected data is another aspect that should be considered. It must be kept in mind that, depending on the selected language, the DI of different guest groups is analyzed. If the DI of tourists of a specific country is planned to be analyzed – like in the research of Liu et al. (2019), and Hu et al. (2014) – it is advisable to choose comments and reviews in the specific language. If a general DI from international guest is aimed to be found the content can be collected in English.

In addition, DI is dependent on travel status, meaning whether the destination has been visited before or not (Jani & Hwang, 2011). The demographics of tourists can also have an impact (Alonso-Almeida et al., 2019). Future online DI analyses should therefore take these aspects into account from the very beginning, to capture an as comprehensive DI as possible. In other words, search and filter criteria should be defined from the very start, according to the strategic goals. Furthermore, it may be interesting for destination managers to link the described experiences with the demographic data of the users, to design and market the destination offers more target group precisely.

Finally, the time span for data collection must be chosen. Again, it depends on the strategic goal whether a snapshot of the current situation is aimed to be analyzed or a development of the DI over the last years. Past studies chose time frames between four month and eleven years.

Collection Phase

After the data selection the collection phase begins. For this stage it must be determined whether the researchers are going to collect the data manually – e.g., with the help of search engines – or if the collection will be performed automatically, with the help of crawling programs like Leximancer or WebLyzard. It is urged to choose automatic

collection, due to the vast amount of online data available nowadays. Manual collection would either be very time consuming or entail the need to reduce the amount of collected data immensely, which leads to a narrower and more inaccurate DI. However, it must be regarded that certain crawling software only works with English dictionaries and therefore cannot be applied in all cases.

In addition, it should be considered whether certain filter criteria shall be taken into account during the data collection. For instance, reviews and posts could be collected sorted by seasons or ratings, to simplify later classification. For the downloading either a web crawler or a web copy program is needed. Along with this, attention should be paid to the recording and storing of the data. Depending on the type of data collected and further processing of the data, different formats are suitable, e.g., PDF, HTML or Datasets (when including text and photos or videos).

Analysis Phase

For the analysis phase it is important to cleanse and prepare the data before starting with the actual analysis, especially when working computer assisted. This includes debugging the collected data, removing stop words (me, and, but, etc.) and identifying composite words (e.g., Sagrada Familia as one word construct, instead of two separate words).

When analyzing the data manually, a codebook must be developed with a clear definition of the search criteria and categories. Codebooks from prior studies can help with the development (Jani & Hwang, 2011; Költringer & Dickinger, 2015; Marine-Roig & Anton Clavé, 2016). Again, manual analysis is not recommended because of the amount of data available and is only suitable in case of a qualitative and explorative analysis.

Both, in manual and in automatic analysis, the analysis techniques must be defined. These could be counting of words / keywords, grouping keywords into topics, definition of relevant categories. For this the programming language Python can be applied, or software like CATPAC, TextSTAT and WordSTAT, which are recommended by researchers in the database (Dickinger & Lalicic, 2016; Garay & Cánoves, 2015; Jani & Hwang, 2011; Mak, 2017). After content analysis further methods can be applied to get a deeper understanding of tourist's behavior in a destination and their sentiments towards a destination. The most common and proven methods are sentiment analysis, co-occurrence analysis, semantic clustering, and correspondence analysis. Especially in

sentiment analysis it is advised, to divide the data into topics and identify the sentiments for topics. When analyzing pictures, a mise en scene analysis with the help of NVivo software is recommended, as done by Mak (2017).

Many authors applied the help of visualization programs to identify the connections between keywords or topics for instance Leximancer or Gephi (Alonso-Almeida et al., 2019; Liu et al., 2019; Oliveira et al., 2019; Tseng et al., 2015). As well as the mapping of posts and keywords, to show the tourist movement within a destination and the “hot spots” (Miah et al., 2017; Van der Zee & Bertocchi, 2018).

Reporting Phase

The previous defined strategic focus of the research, whether it is a theoretical or practical focus, is crucial for the final phase, the reporting of the results. In case of a theoretical focus, emphasis should be put on the methods used, the general tourist behavior insights gained and the comparison of the analyzed media to media from other studies.

In case of a practical focus the results should be presented as action point for destination managers. In this case it is useful to list the main stakeholders and tourist’s groups of the destination and a list of the most popular areas / sights / activities and topics. Moreover, the expressed sentiments should be presented with a clear definition of what is loved and what criticized. As well as the found gap between the perceived and the projected DI.

Following this guideline and taking decisions for all listed points in advance of the actual research, could help to perform a stringent and fluent DI analysis, without jumping back and forth for adjustments to often.

Discussion for future research

However, the described method of online DI analysis offers only a partial section of the DI. The more comprehensive and holistic the DI is to be mapped the more analysis methods should be used. A benchmark analysis with other comparable tourism regions is useful to filter out the unique selling propositions of a destination. For this purpose, collaborations with other destinations could be started.

Furthermore, future online DI analyses could use additional qualitative methods in order to better understand the motives and motivations of tourists and other stakeholders. The

analysis of UGC described above serves to measure and understand the current state. Meaning what emotions, thoughts and evaluations do current and future tourists have about the destination. In order to align the marketing products and strategies around the destination correctly for the future, additional qualitative methods could help. Examples of these methods are focus groups, individual interviews, and observations. The method of DI detection through UGC could serve as a starting point for the development of further qualitative methods. For instance, focus groups could be used to work out concrete solutions for situations that were noted negatively in the online DI. In qualitative individual interviews, representatives of different tourist groups could be interviewed about the identified sentiments in order to gain a deeper understanding of the motives, expectations and motivations for choosing a destination. At the same time, interviews could also be conducted with local stakeholders, such as gastronomes, retailers, and tour guides /attraction providers. They could first be asked about their perceived DI, as well as about the observed behavior of the tourists. In a second step, they could be confronted with the DI drawn online and check to what extent their perception on site deviates from it.

Another possibility for additional qualitative data collection could be the observation of tourist behavior in the destination. Where do they spend a particularly long time? Do they separate from their families or stay together? Do they choose public transportation or cabs at the airport? At what times of day are they traveling and where? Here, again, the insights into tourism flows gained previously in the online DI analysis could be used as a basis. The problems found via UGC analysis can be studied more specifically through observations to find concrete solutions. For example, if it is noted in UGC that public transportation access points at popular attractions are highly crowded, researchers could use observations to analyze tourist flows and provide concrete suggestions for improvement. In this case, for example, better signage to the next subway station to equalize the flow of people.

6 Limitations

This research paper has several limitations. First, this study attempted to provide a holistic view of online DI research between 2010 and 2020, however since research papers were retrieved only from the three databases Web of Science, ProQuest and b-on, contributions on American destinations are missing in the collection (only exception: Deng and Li (2018) research on New York City). Therefore, a research with the same keywords and time settings should be conducted using other databases, which are more specialized on scientific content from America, to complete the picture and compare the findings with the findings of the present study.

Moreover, a dominance of studies on China and Spain were found in the database. It would be interesting to see if this shifts by choosing other databases or if nations like Russians and Germans are generally underrepresented in online DI research. Both can be an interesting subject for future research. Either scientific databases in the field of tourism could be divided in predominant tourism nations, or online DI analysis from the perspective of underrepresented tourists in this database could be conducted.

Additionally, the studies collected have not been assessed on quality and relevance by the author of the present work. Meaning that all contributions in the dataset have been treated equally, no matter how often they have been quoted by others or which ranking they have in a journal ranking list. Therefore, this study's results may include ideas that are not widely supported in tourism research. For future research it is recommended to check the value of collected contributions with the help of a journal ranking list.

When including the research objects into the database a coding matrix was designed to capture the key parameters of each study in a standardized and thus comparable manner. Nevertheless, there is always a risk of over-standardization or loss of important detailed information in this kind of standardization process. Moreover, the coding was conducted by only one researcher which could have led to a bias in capturing the key parameters, since the dual-control principle was not applied. Hence, the lack of detail information and detailed method description limits the present study in its comprehensiveness and therefore is another limitation.

In conclusion, the results of the study should be seen in context with its limitations.

7 Conclusion

*„Aber ein jeder, der in der Ferne ein Land studieren will, er habe es früher nun selbst gesehen oder nicht, wird immer soviel Zeugen aufsuchen, als er nur kann. / But anyone who wants to study a country far away, whether he has seen it himself in the past or not, will always seek out as many observers as he can.“
(Goethe, 1811)*

This wisdom from Goethe applies not only to traveling, but also to all other undertakings. The tourist, when informing himself about his next potential vacation destination, nowadays no longer does this solely with a guidebook, an advertisement of the destination or the report of an acquaintance but is offered almost infinite possibilities of information on the Internet: From reviews, over discussion and question forums, up to detailed travel reports with texts, pictures, and video. From the resulting picture, he can make up his own mind about the region. Tourism marketing organizations and tourism researchers therefore have a great interest in capturing this image created on the Internet as precisely as possible. Therefore, the well-known theory of DI was extended to online DI. The possibility of collecting vast amount of data for the DI analysis requires new methods and tools in data collection and analysis. To not get lost in big data but to use it to its full potential.

Before a destination manager or a researcher starts to conduct a study on the DI online of his region himself, he should refer to Goethe's wisdom and consult as many researchers as possible, who are already experienced in this field of analysis.

The present study was designed to do exactly that. Its aim was to close this research gap and provide a systematic overview of the current situation in the research field of online DI analysis through UGC. The present state of the art was attempted to be explored in an unbiased and holistic way through the method of SLR. The intention was to examine and provide future research paths addressing the DI online.

Therefore, research work on online DI of the last ten years was collected and analyzed to portray the current methodological approach in this scientific field. The outcome is an overview model that can be used as an initial guidance for designing an online DI study according to current standards. Considering the several limitations present in this research it still can be an aid in the development phase of future online DI studies.

References

- Abburu, D. S., & Babu, G. S. (2013). A framework for web information extraction and analysis. *INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY*, 7(2), 574–579. <https://doi.org/10.24297/ijct.v7i2.3459>
- Agarwal, B., & Mittal, N. (2016). Machine Learning Approach for Sentiment Analysis. In B. Agarwal & N. Mittal (Eds.), *Prominent Feature Extraction for Sentiment Analysis* (pp. 21–45). Springer International Publishing. https://doi.org/10.1007/978-3-319-25343-5_3
- airbnb.com. (2021). *About Us*. Airbnb Newsroom. <https://news.airbnb.com/about-us/>
- Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*, 58(2), 175–191. <https://doi.org/10.1177/0047287517747753>
- Alonso-Almeida, Borrajo-Millán, & Yi. (2019). Are Social Media Data Pushing Overtourism? The Case of Barcelona and Chinese Tourists. *Sustainability*, 11(12), 3356. <https://doi.org/10.3390/su11123356>
- Alrawadieh, Z., Dincer, M. Z., Istanbulu Dincer, F., & Mammadova, P. (2018). Understanding destination image from the perspective of Western travel bloggers: The case of Istanbul. *International Journal of Culture, Tourism and Hospitality Research*, 12(2), 198–212. <https://doi.org/10.1108/IJCTHR-12-2017-0124>
- Ariyasriwatana, W., & Quiroga, L. (2016). A Thousand Ways to Say 'Delicious!'- Categorizing Expressions of Deliciousness from Restaurant Reviews on the Social Network Site Yelp. *Appetite*, 104. <https://doi.org/10.1016/j.appet.2016.01.002>
- Armstrong, G., Adam, S., Denize, S., & Kotler, P. (2014). *Principles of Marketing*. Pearson Australia.

- Assaker, G. (2014). Examining a hierarchical model of Australia's destination image. *Journal of Vacation Marketing*, 20(3), 195–210. <https://doi.org/10.1177/1356766714527104>
- Augsten, S. (2019, July 12). *Was ist Python?* Dev Insider. <https://www.dev-insider.de/was-ist-python-a-843060/>
- Baloglu, S., & McCleary, K. W. (1999). A model of destination image formation. *Annals of Tourism Research*, 26(4), 868–897. [https://doi.org/10.1016/S0160-7383\(99\)00030-4](https://doi.org/10.1016/S0160-7383(99)00030-4)
- Basaran, U. (2016). Examining the Relationships of Cognitive, Affective, and Conative Destination Image: A Research on Safranbolu, Turkey. *International Business Research*, 9(5), 164. <https://doi.org/10.5539/ibr.v9n5p164>
- Behdenna, S., Barigou, F., & Belalem, G. (2018). Document Level Sentiment Analysis: A survey. *EAI Endorsed Transactions on Context-Aware Systems and Applications*, 4, 154339. <https://doi.org/10.4108/eai.14-3-2018.154339>
- Belarmino, A., & Koh, Y. (2020). A critical review of research regarding peer-to-peer accommodations. *International Journal of Hospitality Management*, 84, 102315. <https://doi.org/10.1016/j.ijhm.2019.05.011>
- Bem, D. J. (1995). Writing a review article for Psychological Bulletin. *Psychological Bulletin*, 118(2), 172–177. <https://doi.org/10.1037/0033-2909.118.2.172>
- b-on. (2020, June). *What is b-on?* <https://www.b-on.pt/en/what-is-b-on/>
- Brosius, H.-B., Koschel, F., & Haas, A. (Eds.). (2009). Inhaltsanalyse II: Kategorien und Codebuch. In *Methoden der empirischen Kommunikationsforschung: Eine Einführung* (pp. 154–171). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91821-1_9

-
- Brownlee, J. (2020, April 5). 10 Clustering Algorithms With Python. *Machine Learning Mastery*. <https://machinelearningmastery.com/clustering-algorithms-with-python/>
- Bucur, C. (2015). Using Opinion Mining Techniques in Tourism. *Procedia Economics and Finance*, 23, 1666–1673. [https://doi.org/10.1016/S2212-5671\(15\)00471-2](https://doi.org/10.1016/S2212-5671(15)00471-2)
- Buultjens, J., Neale, K., & Lamont, M. (2013). Hosts, guests and a drug culture in a destination: A case study of Nimbin, Australia. *Journal of Destination Marketing & Management*, 2(3), 185–195. <https://doi.org/10.1016/j.jdmm.2013.08.003>
- Buzydlowski, J. W. (2015). *Co-occurrence analysis as a framework for data mining*. 6, 19.
- Calantone, R. J., Di Benedetto, C. A., Hakam, A., & Bojanic, D. C. (1989). Multiple Multinational Tourism Positioning Using Correspondence Analysis. *Journal of Travel Research*, 28(2), 25–32. <https://doi.org/10.1177/004728758902800207>
- Cao, M., Liang, J., Li, M., Zhou, Z., & Zhu, M. (2020). TDIVis: Visual analysis of tourism destination images. *Frontiers of Information Technology & Electronic Engineering*, 21(4), 536–557. <https://doi.org/10.1631/FITEE.1900631>
- Carter, R. W. (Bill), Thok, S., O'Rourke, V., & Pearce, T. (2015). Sustainable tourism and its use as a development strategy in Cambodia: A systematic literature review. *Journal of Sustainable Tourism*, 23(5), 797–818. <https://doi.org/10.1080/09669582.2014.978787>
- Chaffey, Dave. (2020, April 17). *Global social media research summary 2020*. Smart Insights. <https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/>

- Chen, C.-F., & Tsai, D. (2007). How destination image and evaluative factors affect behavioral intentions? *Tourism Management*, 28(4), 1115–1122. <https://doi.org/10.1016/j.tourman.2006.07.007>
- Chen, Y., Liu, Y., Wang, C., Chen, K. (2019). *Is hosting tourism events a sustainable way of destination branding: Evidences from Sanya, China.* <https://doi.org/10.1080/23311886.2019.1692456>
- Chi, C. G.-Q., & Qu, H. (2008). Examining the structural relationships of destination image, tourist satisfaction and destination loyalty: An integrated approach. *Tourism Management*, 29(4), 624–636. <https://doi.org/10.1016/j.tourman.2007.06.007>
- Chiu, C., Chiu, N.-H., Sung, R.-J., & Hsieh, P.-Y. (2015). Opinion mining of hotel customer-generated contents in Chinese weblogs. *Current Issues in Tourism*, 18(5), 477–495. <https://doi.org/10.1080/13683500.2013.841656>
- Choi, S., Lehto, X. Y., & Morrison, A. M. (2007). Destination image representation on the web: Content analysis of Macau travel related websites. *Tourism Management*, 28(1), 118–129. <https://doi.org/10.1016/j.tourman.2006.03.002>
- Clarke, H., & Hassanien, A. (2020). An Evaluation of Toronto’s Destination Image Through Tourist Generated Content on Twitter: *International Journal of Customer Relationship Marketing and Management*, 11(2), 1–16. <https://doi.org/10.4018/IJCRMM.2020040101>
- Correia, A., Oom do Valle, P., & Moço, C. (2007). Why people travel to exotic places. *International Journal of Culture, Tourism and Hospitality Research*, 1(1), 45–61. <https://doi.org/10.1108/17506180710729600>

- Crompton, J. L. (1979). An Assessment of the Image of Mexico as a Vacation Destination and the Influence of Geographical Location Upon That Image. *Journal of Travel Research*, 17(4), 18–23. <https://doi.org/10.1177/004728757901700404>
- Degen, H. (1999). Menschen, Produkte, Märkte. In H. Degen (Ed.), *Kundenorientierte Softwareproduktion* (pp. 69–229). Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-663-08679-6_2
- Deng, N., & Li, X. (Robert). (2018). Feeling a destination through the “right” photos: A machine learning model for DMOs’ photo selection. *Tourism Management*, 65, 267–278. <https://doi.org/10.1016/j.tourman.2017.09.010>
- Dickinger, A., & Lalicic, L. (2016). An analysis of destination brand personality and emotions: A comparison study. *Information Technology & Tourism*, 15(4), 317–340. <https://doi.org/10.1007/s40558-015-0044-x>
- Donnelly, R., & Manolova, T. S. (2020). Foreign location decisions through an institutional lens: A systematic review and future research agenda. *International Business Review*, 29(4), 101690. <https://doi.org/10.1016/j.ibusrev.2020.101690>
- Doyle, C. (2011). Pollyanna effect. In *A Dictionary of Marketing*. Oxford University Press. <https://www.oxfordreference.com/view/10.1093/acref/9780199590230.001.0001/acref-9780199590230-e-1348>
- Duman, T., Ozbal, O., & Duerod, M. (2018). The role of affective factors on brand resonance: Measuring customer-based brand equity for the Sarajevo brand. *Journal of Destination Marketing & Management*, 8, 359–372. <https://doi.org/10.1016/j.jdmm.2017.08.001>
- Echtner, C. M., & Ritchie, J. R. B. (1991). The meaning and measurement of destination image. *Journal of Tourism Studies*, 2(2), 2–12.

- Echtner, C. M., & Ritchie, J. R. B. (1993). The Measurement of Destination Image: An Empirical Assessment. *Journal of Travel Research*, 31(4), 3–13. <https://doi.org/10.1177/004728759303100402>
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.
- Ficamos, P., & Liu, Y. (2016). A Topic based Approach for Sentiment Analysis on Twitter Data. *International Journal of Advanced Computer Science and Applications*, 7. <https://doi.org/10.14569/IJACSA.2016.071226>
- Figuroa-Domecq, C. (2015). Tourism gender research: A critical accounting. *Annals of Tourism Research*, 17.
- Fisch, C., & Block, J. (2018). Six tips for your (systematic) literature review in business and management research. *Management Review Quarterly*, 68(2), 103–106. <https://doi.org/10.1007/s11301-018-0142-x>
- Gálvez-Rodríguez, M. del M., Alonso-Cañadas, J., Haro-de-Rosario, A., & Caba-Pérez, C. (2020). Exploring best practices for online engagement via Facebook with local destination management organisations (DMOs) in Europe: A longitudinal analysis. *Tourism Management Perspectives*, 34, 100636. <https://doi.org/10.1016/j.tmp.2020.100636>
- Garay, L. (2019). #Visitspain. Breaking down affective and cognitive attributes in the social media construction of the tourist destination image. *Tourism Management Perspectives*, 32, 100560. <https://doi.org/10.1016/j.tmp.2019.100560>
- Garay, L., & Cànoves, G. (2015). Barcelona seen through the eyes of TripAdvisor: Actors, typologies and components of destination image in social media platforms. *Current Issues in Tourism*, 20(1), 33–37. <https://doi.org/10.1080/13683500.2015.1073229>

- Garay, L., & Morales Pérez, S. (2017). Understanding the creation of destination images through a festival's Twitter conversation. *International Journal of Event and Festival Management*, 8(1), 39–54. <https://doi.org/10.1108/IJEFM-04-2016-0030>
- Gartner, W. C. (1986). Temporal influences on image change. *Annals of Tourism Research*, 13(4), 635–644. [https://doi.org/10.1016/0160-7383\(86\)90006-X](https://doi.org/10.1016/0160-7383(86)90006-X)
- Gartner, W. C. (1994). Image Formation Process. *Journal of Travel & Tourism Marketing*, 2(2–3), 191–216. https://doi.org/10.1300/J073v02n02_12
- Gephi—The Open Graph Viz Platform. (n.d.). Retrieved May 22, 2021, from <https://gephi.org/>
- Gkritzali, A., Gritzalis, D., & Stavrou, V. (2018). Is Xenios Zeus Still Alive? Destination Image of Athens in the Years of Recession. *Journal of Travel Research*, 57(4), 540–554. <https://doi.org/10.1177/0047287517705225>
- Goethe, J. W. (1796). *Wilhelm Meisters Lehrjahre* (Ehrhard Bahr). Reclam, Philipp. https://www.thalia.de/shop/home/artikeldetails/ID02894782.html?ProvID=11000522&gclid=Cj0KCQjwna2FBhDPArisACAEC_VkXblioIHTgEC_Ko5bXXb7wM4DX48GeYEXaw1p_ZQ54vioCl6ENPwaAln0EALw_wcB
- Goethe, J. W. (1811). *Theoretische Schriften*. Cottaische Buchhandlung.
- González-Rodríguez, M. R., Martínez-Torres, R., & Toral, S. (2016). Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. *International Journal of Contemporary Hospitality Management*, 28(11), 2609–2627. <https://doi.org/10.1108/IJCHM-02-2015-0057>
- Govers, R. (2012). Brand Dubai and its competitors in the Middle East: An image and reputation analysis. *Place Branding and Public Diplomacy*, 8(1), 48–57. <https://doi.org/10.1057/pb.2011.30>

- Güneren, E., Tepavcevic, J., & Pektas, F. (2019). The impact of Destination Image on Destination Recommendation. *International Journal of Management Economics and Business*, 15. <https://doi.org/10.17130/ijmeb.2019456406>
- Gunn, C. A. (1988). Vacationscape: Designing tourist regions. *Vacationscape: Designing Tourist Regions*. <https://www.cabdirect.org/cabdirect/abstract/19901880817>
- Herkommer, C. (2012). Die computergestützte qualitative Inhaltsanalyse. Eine Möglichkeit zur Erweiterung des Methodenkanons der (zeit-) historischen Forschung. *Zeitschrift für digitale Geschichtswissenschaften*, 1, 17.
- Hu, T., Marchiori, E., Kalbaska, N., & Cantoni, L. (2014). Online representation of Switzerland as a tourism destination: An exploratory research on a Chinese microblogging platform. *Studies in Communication Sciences*, 14(2), 136–143. <https://doi.org/10.1016/j.scoms.2014.12.001>
- Hunt, J. D. (1975). Image as a Factor in Tourism Development. *Journal of Travel Research*, 13(3), 1–7. <https://doi.org/10.1177/004728757501300301>
- Iglesias-Sánchez, P. P., Correia, M. B., Jambrino, C., & De Las Heras-Pedrosa, C. (2020). Instagram as a Co-Creation Space for Tourist Destination Image-Building: Algarve and Costa del Sol Case Studies. *Sustainability*, 12, 2793. <https://doi.org/10.3390/su12072793>
- Instagram.com (2021). Retrieved July 7, 2021, from <https://www.instagram.com/>.
- Jalilvand, M. R., & Heidari, A. (2017). Comparing face-to-face and electronic word-of-mouth in destination image formation. *Information Technology & People; West Linn*, 30(4), 710–735. <http://dx.doi.org/10.1108/ITP-09-2016-0204>
- Jani, D., & Hwang, Y.-H. (2011). User-generated Destination Image through Weblogs: A Comparison of Pre- and Post-visit Images. *Asia Pacific Journal of Tourism Research*, 16(3), 339–356. <https://doi.org/10.1080/10941665.2011.572670>

- Keller, K. L. (1993). Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing*, 57(1), 1–22.
<https://doi.org/10.1177/002224299305700101>
- Khoo, C. S., & Johnkhan, S. B. (2018). Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *Journal of Information Science*, 44(4), 491–511. <https://doi.org/10.1177/0165551517703514>
- Kim, J.-H. (2014). The antecedents of memorable tourism experiences: The development of a scale to measure the destination attributes associated with memorable experiences. *Tourism Management*, 44, 34–45.
<https://doi.org/10.1016/j.tourman.2014.02.007>
- Kim, K., Park, O., Yun, S., & Yun, H. (2017). What makes tourists feel negatively about tourism destinations? Application of hybrid text mining methodology to smart destination management. *Technological Forecasting and Social Change*, 123, 362–369. <https://doi.org/10.1016/j.techfore.2017.01.001>
- Kim, S. S., Lee, C.-K., & Klenosky, D. B. (2003). The influence of push and pull factors at Korean national parks. *Tourism Management*, 24(2), 169–180.
[https://doi.org/10.1016/S0261-5177\(02\)00059-6](https://doi.org/10.1016/S0261-5177(02)00059-6)
- Kirilenko, A. P., & Stepchenkova, S. O. (2017). Sochi 2014 Olympics on Twitter: Perspectives of hosts and guests. *Tourism Management*, 63, 54–65.
<https://doi.org/10.1016/j.tourman.2017.06.007>
- Koch, K. (2020, March 26). *A Friendly Introduction to Text Clustering*. Medium.
<https://towardsdatascience.com/a-friendly-introduction-to-text-clustering-fa996bcefd04>

- Költringer, C., & Dickinger, A. (2015). Analyzing destination branding and image from online sources: A web content mining approach. *Journal of Business Research*, 68(9), 1836–1843. <https://doi.org/10.1016/j.jbusres.2015.01.011>
- Kuhn, A., Ducasse, S., & Gîrba, T. (2007). Semantic clustering: Identifying topics in source code. *Information and Software Technology*, 49(3), 230–243. <https://doi.org/10.1016/j.infsof.2006.10.017>
- Law, R., & Cheung, S. (2010). The Perceived Destination Image of Hong Kong as Revealed in the Travel Blogs of Mainland Chinese Tourists. *International Journal of Hospitality & Tourism Administration*, 11(4), 303–327. <https://doi.org/10.1080/15256480.2010.518521>
- Leung, D., Law, R., & Lee, H. A. (2010). The perceived destination image of Hong Kong on Ctrip.com. *International Journal of Tourism Research*, n/a-n/a. <https://doi.org/10.1002/jtr.803>
- Li, J. (Justin), Ali, F., & Kim, W. G. (2015). Reexamination of the role of destination image in tourism: An updated literature review. *E-Review of Tourism Research*, Vol. 12 Issue 3/4, 191–209.
- Li, Y. R., Lin, Y. C., Tsai, P. H., & Wang, Y. Y. (2015). Traveller-Generated Contents for Destination Image Formation: Mainland China Travellers to Taiwan as a Case Study. *Journal of Travel & Tourism Marketing*, 32(5), 518–533. <https://doi.org/10.1080/10548408.2014.918924>
- Lian, T., & Yu, C. (2019). Impacts of online images of a tourist destination on tourist travel decision. *Tourism Geographies*, 21(4), 635–664. <https://doi.org/10.1080/14616688.2019.1571094>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA

- statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ*, 339, b2700. <https://doi.org/10.1136/bmj.b2700>
- Lim, W. M., Yap, S.-F., & Makkar, M. (2021). Home sharing in marketing and tourism at a tipping point: What do we know, how do we know, and where should we be heading? *Journal of Business Research*, 122, 534–566. <https://doi.org/10.1016/j.jbusres.2020.08.051>
- Lin, Y.-S., & Huang, J.-Y. (2006). Internet blogs as a tourism marketing medium: A case study. *Journal of Business Research*, 59(10), 1201–1205. <https://doi.org/10.1016/j.jbusres.2005.11.005>
- Liu, B. (2015). *Sentiment Analysis* (1. Edition). Cambridge University Press.
- Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management*, 71, 337–347. <https://doi.org/10.1016/j.tourman.2018.10.004>
- Lopes, S. D. F. (2011). Destination image: Origins, Developments and Implications. *PASOS. Revista de Turismo y Patrimonio Cultural*, 9(2), 305–315. <https://doi.org/10.25145/j.pasos.2011.09.027>
- Lopez-Cordova, E. (2020). *Digital Platforms and the Demand for International Tourism Services*. The World Bank. <https://doi.org/10.1596/1813-9450-9147>
- Madden, K., Rashid, B., & Zainol, N. A. (2016). Beyond the motivation theory of destination image. *Tourism and Hospitality Management*, 22(2), 247–264. <https://doi.org/10.20867/thm.22.2.1>

- Mair, J., Ritchie, B. W., & Walters, G. (2014). *Towards a research agenda for post-disaster and post-crisis recovery strategies for tourist destinations: A narrative review*. 28.
- Mak, A. H. N. (2017). Online destination image: Comparing national tourism organisation's and tourists' perspectives. *Tourism Management*, 60, 280–297. <https://doi.org/10.1016/j.tourman.2016.12.012>
- Marine-Roig, E. (2017). Measuring Destination Image through Travel Reviews in Search Engines. *Sustainability*, 9(8), 1425. <https://doi.org/10.3390/su9081425>
- Marine-Roig, E. (2019). Destination Image Analytics Through Traveller-Generated Content. *Sustainability*, 11(12), 3392. <https://doi.org/10.3390/su11123392>
- Marine-Roig, E., & Anton Clavé, S. (2016). A detailed method for destination image analysis using user-generated content. *Information Technology & Tourism*, 15(4), 341–364. <https://doi.org/10.1007/s40558-015-0040-1>
- Marine-Roig, E., & Ferrer-Rosell, B. (2018). Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis. *Tourism Management*, 68, 236–249. <https://doi.org/10.1016/j.tourman.2018.03.020>
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370.
- Mayring, P. (2004). Qualitative Content Analysis. *A Companion to Qualitative Research*, 1(2), 159–176.
- Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A Big Data Analytics Method for Tourist Behaviour Analysis. *Information & Management*, 54(6), 771–785. <https://doi.org/10.1016/j.im.2016.11.011>
- Micera, R., & Crispino, R. (2017). Destination web reputation as “smart tool” for image building: The case analysis of Naples city-destination. *International Journal of*

- Tourism Cities; Bingley*, 3(4), 406–423. <http://dx.doi.org/10.1108/IJTC-11-2016-0048>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Munzero, M., Montero, C. S., Sutinen, E., & Pajunen, J. (2014). Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text. *IEEE Transactions on Affective Computing*, 5(2), 101–111. <https://doi.org/10.1109/TAFFC.2014.2317187>
- NCapture*. (n.d.). Retrieved May 22, 2021, from <https://help-nv.qsrinternational.com/20/win/Content/ncapture/ncapture.htm>
- Nielsen, B. B., Asmussen, C. G., & Weatherall, C. D. (2017). The location choice of foreign direct investments: Empirical evidence and methodological challenges. *Journal of World Business*, 52(1), 62–82. <https://doi.org/10.1016/j.jwb.2016.10.006>
- Oliveira, C., Brochado, A., Moro, S., & Paulo, R. (2019). *Consumer perception of tourist experience through online reviews*. 696–717. <https://doi.org/10.1108/WHATT-09-2019-0052>
- Oliver, R. L., & Winer, R. S. (1987). A framework for the formation and structure of consumer expectations: Review and propositions. *Journal of Economic Psychology*, 8(4), 469–499. [https://doi.org/10.1016/0167-4870\(87\)90037-7](https://doi.org/10.1016/0167-4870(87)90037-7)
- Otto, J. E., & Ritchie, J. R. B. (1996). The service experience in tourism. *Tourism Management*, 17(3), 165–174. [https://doi.org/10.1016/0261-5177\(96\)00003-9](https://doi.org/10.1016/0261-5177(96)00003-9)

- Palmer, S. (2020, September 7). *LibGuides: ProQuest Central: About*. Proquest.Libguides. <https://proquest.libguides.com/pqc/about>
- Paul, J., & Criado, A. R. (2020). The art of writing literature review: What do we know and what do we need to know? *International Business Review*, 29(4), 101717. <https://doi.org/10.1016/j.ibusrev.2020.101717>
- Picazo, P., Moreno-Gil, S., & Gursoy, D. (2019). Analysis of the projected image of tourism accommodations: A methodological proposal. *International Journal of Contemporary Hospitality Management*, 31(8), 3325–3351. <https://doi.org/10.1108/IJCHM-05-2017-0276>
- Pickering, C., & Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research & Development*, 33(3), 534–548. <https://doi.org/10.1080/07294360.2013.841651>
- Pike, S., & Ryan, C. (2004). Destination Positioning Analysis through a Comparison of Cognitive, Affective, and Conative Perceptions. *Journal of Travel Research*, 42(4), 333–342. <https://doi.org/10.1177/0047287504263029>
- Rahmani, K., Gnoth, J., & Mather, D. (2019). A Psycholinguistic View of Tourists' Emotional Experiences. *Journal of Travel Research*, 58(2), 192.
- Ramsenthaler, C. (2013). Was ist „Qualitative Inhaltsanalyse?“. In M. Schnell, C. Schulz, H. Kolbe, & C. Dunger (Eds.), *Der Patient am Lebensende: Eine Qualitative Inhaltsanalyse* (pp. 23–42). Springer Fachmedien. https://doi.org/10.1007/978-3-531-19660-2_2
- Ren, G., & Hong, T. (2017). Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach. *Sustainability*, 9, 1765. <https://doi.org/10.3390/su9101765>

- Robino, D. M. (2019). *Global Destination Cities Index* (p. 16). Mastercard.
<https://newsroom.mastercard.com/wp-content/uploads/2019/09/GDCI-Global-Report-FINAL-1.pdf>
- Roemer, E. (2014). Internationales Marketing Management. In *Internationales Marketing Management* (pp. 43–86). Schäffer-Poeschel Verlag.
<https://doi.org/10.34156/9783799267441-43>
- Rössler, P. (2017). *Inhaltsanalyse*. UTB.
- San Martín, H., & Rodríguez del Bosque, I. A. (2008). Exploring the cognitive–affective nature of destination image and the role of psychological factors in its formation. *Tourism Management*, 29(2), 263–277.
<https://doi.org/10.1016/j.tourman.2007.03.012>
- Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *Journal of Travel & Tourism Marketing*, 32(5), 608–621. <https://doi.org/10.1080/10548408.2014.933154>
- Sharma, A., Park, S., & Nicolau, J. L. (2020). Testing loss aversion and diminishing sensitivity in review sentiment. *Tourism Management*, 77, 104020.
<https://doi.org/10.1016/j.tourman.2019.104020>
- Song, Z., Su, X., & Liaoning Li. (2013). The Indirect Effects of Destination Image on Destination Loyalty Intention Through Tourist Satisfaction and Perceived Value: The Bootstrap Approach. *Journal of Travel & Tourism Marketing*, 30(4), 386–409. <https://doi.org/10.1080/10548408.2013.784157>
- Sönmez, S., & Sirakaya, E. (2002). A Distorted Destination Image? The Case of Turkey. *Journal of Travel Research*, 41(2), 185–196.
<https://doi.org/10.1177/004728702237418>

- Sun, G. (2020). Symmetry Analysis in Analyzing Cognitive and Emotional Attitudes for Tourism Consumers by Applying Artificial Intelligence Python Technology. *Symmetry*, *12*(4), 606. <https://doi.org/10.3390/sym12040606>
- Sun, M., Ryan, C., & Pan, S. (2015). Using Chinese Travel Blogs to Examine Perceived Destination Image: The Case of New Zealand. *Journal of Travel Research*, *54*(4), 543–555. <https://doi.org/10.1177/0047287514522882>
- Tasci, A. D. A., Gartner, W. C., & Cavusgil, S. T. (2007). Measurement of destination brand bias using a quasi-experimental design. *Tourism Management*, *28*(6), 1529–1540. <https://doi.org/10.1016/j.tourman.2007.02.009>
- Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*(1), 163–173. <https://doi.org/10.1002/asi.21662>
- Tkachuk, L. (2018). *The current trends of aviation and tourism cooperation management in condition of geopolitical uncertainty*.
- Toral, S. L., Martínez-Torres, M. R., & Gonzalez-Rodriguez, M. R. (2018). Identification of the Unique Attributes of Tourist Destinations from Online Reviews. *Journal of Travel Research*, *57*(7), 908–919. <https://doi.org/10.1177/0047287517724918>
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, *14*(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Trommsdorff, G. (1989). Sozialisation und Werthaltungen im Kulturvergleich. In G. Trommsdorff (Ed.), *Sozialisation im Kulturvergleich* (pp. 97–121). De Gruyter. <https://doi.org/10.1515/9783110511147-007>

- Tseng, C., Wu, B., Morrison, A. M., Zhang, J., & Chen, Y. (2015). Travel blogs on China as a destination image formation agent: A qualitative analysis using Leximancer. *Tourism Management*, 46, 347–358. <https://doi.org/10.1016/j.tourman.2014.07.012>
- Vada, S., Prentice, C., Scott, N., & Hsiao, A. (2020). Positive psychology and tourist well-being: A systematic literature review. *Tourism Management Perspectives*, 33, 100631. <https://doi.org/10.1016/j.tmp.2019.100631>
- Van der Zee, E., & Bertocchi, D. (2018). Finding patterns in urban tourist behaviour: A social network analysis approach based on TripAdvisor reviews. *Information Technology & Tourism*, 20(1–4), 153–180. <https://doi.org/10.1007/s40558-018-0128-5>
- Watson, R., Wilson, H. N., Smart, P., & Macdonald, E. K. (2018). Harnessing Difference: A Capability-Based Framework for Stakeholder Engagement in Environmental Innovation. *Journal of Product Innovation Management*, 35(2), 254–279. <https://doi.org/10.1111/jpim.12394>
- Weber, K. (1997). The assessment of tourist satisfaction using the expectancy disconfirmation theory: A study of the German travel market in Australia. *Pacific Tourism Review*, 1(1), 35–45.
- Yang, E. C. L., Khoo-Lattimore, C., & Arcodia, C. (2017). A systematic literature review of risk and gender research in tourism. *Tourism Management*, 58, 89–100. <https://doi.org/10.1016/j.tourman.2016.10.011>
- Yu, C.-E., & Zhang, X. (2020). The embedded feelings in local gastronomy: A sentiment analysis of online reviews. *Journal of Hospitality and Tourism Technology*, 11(3), 461–478. <https://doi.org/10.1108/JHTT-02-2019-0028>

Appendix

Appendix 1: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 1)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
The Perceived Destination Image of Hong Kong on Ctrip.com	International Journal of Tourism Research	Daniel Leung Rob Law Hee Andy Lee	2010	China	Tourism management	photo metadata analysis	Frequency Analysis of positive and negative perceptions		Ctrip.com	n= 2,247 blog entries t= 7 years	-	-	Practical: DI of Hong Kong as Shopping Paradise, advise for DMO.
The Perceived Destination Image of Hong Kong as Revealed in the Travel Blogs of Mainland Chinese Tourists	International Journal of Hospitality & Tourism Administration	Rob Law Shannon Cheung	2010	China	Tourism management	Sentiment Analysis	Frequency Analysis of positive and negative perceptions		Travelblogs	n= 120 blog entries t= 4 months	-	-	Practical: DI of Hong Kong by Chinese tourists
User-generated Destination Image through Weblogs: A Comparison of Pre- and Post-visit Images	Asia Pacific Journal of Tourism Research	Dev Jani Yeong-Hyeon Hwang	2011	Tanzania	Tourism management	Topic based sentiment analysis	-	Text STAT software	Lonely Planet Forum Posts	n= 214 posts t= 1 year	-	four main destination attributes: accommodation, environment, activities, infrastructure	Theoretical: descriptors that reflect psychological attributes are used more often by tourists who have already visited the island. Tourists who have already visited a place are more reflective about culture and nature than non-visitors. Practical: DI of Zanzibar. Advice for narrowing the gap between pre- and post-travel DI
Brand Dubai and its competitors in the Middle East: An image and reputation analysis	Place Branding and Public Diplomacy	Robert Govers	2012	UAE	Tourism management	Visual Content Analysis Content Analysis	Automated Sentiment Analysis	Radian6	650 million sources including Twitter, Facebook, YouTube, blogs, news sites and forums	t= 3 years	-	-	Practical: Events lead to a higher number of postings (positive and negative). Image of Dubai and the Middle East.

Appendix 2: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 2)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
Hosts, guests and a drug culture in a destination: A case study of Nimbin, Australia	Journal of Destination Marketing & Management	Jeremy Buultjens, Kate Neale, Matthew Lamont	2013	Australia	Tourism management	Qualitative Content Analysis	Textual analysis	-	Travelblogs, Advisor websites, DMO websites	n= content from 42 websites	-	-	Theoretical: Huge differences between UGC, media content and DMO content Practical: insights about sentiment towards Nimbin.
Online representation of Switzerland as a tourism destination: An exploratory research on a Chinese microblogging platform	Studies in Communication Sciences	Tu Hu, Elena Marchiori, Nadzeya Kalbaska, Lorenzo Cantoni	2014	Switzerland	Tourism management	Content classification analysis. (Content + Sentiment analysis)	-	-	Sina Weibo	n= 732 posts t= 1 year	-	Four new categories: Travel phenomena & stories, Willingness to go, Effective official communication, Link sharing.	Practical: DI of Switzerland perceived by Chinese Tourists
Traveller-Generated Contents for Destination Image Formation: Mainland China Travellers to Taiwan as a Case Study	Journal of Travel & Tourism Marketing	Yan Ru Li, Yao Chin Lin, Ping Heng Tsai, Yung Yueh Wang	2015	Taiwan	Tourism management	Text mining	Content Analysis. Correlation analysis. Cognitive affective components	-	Ctrip.com	n= 1,033 blog articles about Taiwan t= 4 years	-	-	Practical: DI with new insight into the leisure and cultural activities of Hong Kong and Taiwan.
Using Chinese Travel Blogs to Examine Perceived Destination Image: The Case of New Zealand	Journal of Travel Research	Minghui Sun, Chris Ryan, Steve Pan	2015	New Zealand,	Tourism management	Qualitative and quantitative Content Analysis	-	* Computer-assisted qualitative data analysis software (CAQDAS) * SPSS * Atlas ti (for qualitative content analysis)	Chinese travel blogs. Eg: Sina blog, Sohu blog	n= 409 blog entries t= 6 months	-	-	Practical: DI of New Zealand by Chinese tourists, with practical implementation advice.

Appendix 3: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 3)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
Barcelona seen through the eyes of TripAdvisor: actors, typologies and components of destination image in social media platforms	Current Issues in Tourism, 2015	Lluís Garay Tamajón, Gemma Cànoves Valiente	2015	Spain	Tourism management	Content Analysis	-	CATA Software (Computer-Aided Text Analysis)	TripAdvisor.com barcelonatourism.com	n= 949 TripAdvisor posts about Barcelona	-	-	Theoretical: DMO information more exhaustive. UGC more detailed on products. Questions of potential tourists have a more cognitive nature. Answers deliver a mix of cognitive and affective image.
Analyzing destination branding and image from online sources: A web content mining approach	Journal of Business Research	Clemens Költringer, Astrid Dickinger	2015	Austria	Marketing	Automated Content Analysis	Correspondence analysis. Co-occurrence analysis. Sentiment detection.	webLyzard crawling agent (for document collection)	International online travel communities and social travel guides such as www.travelblogs.com. AngloAmerican news media websites, DMO websites.	n= 10,804 online articles about Vienna	-	Sights, Tourism, Activity, Experience	Theoretical: UGC language differs significantly from Media or DMO samples. UGC is the most rich and diverse source. Practical: Destination managers should focus on promoting the destination by employing a more holistic and emotionally connotative communication strategy. Insights about the DI of Vienna
An Analysis of Destination Brand Personality and Emotions: A Comparison Study	Information Technology & Tourism	Astrid Dickinger, Lidija Lalicic	2016	Austria	Marketing	Quantitative Content Analysis	-	WordStat (computer-aided content analysis program)	TripAdvisor.com	n= 1,104 reviews	-	-	Theoretical: survey-based data and review-based data are distinctly different when analyzing emotion-based experiences. Practical: For restaurant / accommodation managers to manage feelings

Appendix 4: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 4)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
A detailed method for destination image analysis using user-generated content	Information Technology & Tourism	Estela Marine-Roig, Salvador Anton Clave	2016	Spain	Tourism management	Qualitative Content Analysis	Positioning of Keywords & Phrases in the Text	-	TravelBlog.org, TravelPod.com, TripAdvisor.com, VirtualTourist.com	n= 132,502 travel blogs and OTRs t= 11years	-	-	Theoretical: Proposes a detailed model for content analysis of large amount of data, with significant insights concerning both the cognitive and the affective components of destination image
Perceived image specialisation in multiscalar tourism destinations	Journal of Destination Marketing & Management	Estela Marine-Roig, Salvador Anton Clavé	2016	Spain	Tourism management	Quantitative Content Analysis	spatial analysis	Language Detection program (to detect the language of the blog entries). Site Content Analyzer software (to conduct keyword count)	TripAdvisor.com, TravelBlog.org, TravelPod.com, VirtualTourist.com	n= 127,895 files	-	Eight categories of attraction factors: 1. Food and Wine, 2. Intangible Heritage, 3. Leisure and Recreation, 4. Nature and Active Tourism, 5. Sports, 6. Sun, Sea and Sand, 7. Tangible Heritage, 8. Urban Environment	Theoretical: The application of spatial indices to the study of tourists' images online has proved to be a useful method to assess brand image specialization in multiscalar destinations. Practical: DI of Catalonia, especially Barcelona. Recommendation to consider geographical scales both when analyzing and building the image of a destination.
Destination web reputation as “smart tool” for image building: the case analysis of Naples city destination	International Journal of Tourism Cities	Roberto Micera, Raffaele Crispino	2017	Italy	Tourism management	Quantitative Content Analysis	Sentiment analysis	Blogmeter (web-based social listening platform)	All social media platforms available through Blogmeter	n= approx. 140,000 internet conversations t= 4 months	-	-	Practical: DI of Naples. Most discussed topics about Naples on Social Media, in English and Italian. Positive and Negative sentiments about Naples. Destinations must inform, tell, converse, and listen

Appendix 5: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 5)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
Measuring Destination Image through Travel Reviews in Search Engines	Sustainability	Estela Marine-Roig	2017	France	Tourism management	Quantitative Content Analysis	Perceived and transmitted destination image	UltraEdit (for extracting meta data)	TripAdvisor.com	n= 387,414 reviews	method design to measure the perceived (and transmitted) TDI from OTRs as presented in search engines.	-	Practical: Managerial Implications for the Tourist region of Paris.
A big data analytics method for tourist behaviour analysis	Information & Management	Shah J. Miah, Huy Quan Vu, John Gammack, Michael McGrath	2017	Australia	Tourism management	photo metadata analysis	geographical data clustering	* General Architecture for Text Engineering (GATE) (for text processing) * Speeded-Up Robust Features (SURF) (feature descriptor for photo content)	Flickr	n= 238,290 photos t= 5 years	the study has developed an IT artefact in the form of a general method for generating meaningful information and predictive insights from geotagged photos.	-	Some practical advice for Melbourne
What makes tourists feel negatively about tourism destinations? Application of hybrid text mining methodology to smart destination management	Technological Forecasting & Social Change	Kun Kim, Oun-joung Park, Seunghyun Yun, Haejung Yun	2017	France	Tourism management	Sentiment Analysis	Co-occurrence analysis	* Python3.3 (programming language) * Stanford sentiment analysis tool based on JAVA1.7.0_65.	virtualltourist.com	n= 19,835 reviews	Hybrid method model: Analysis of travelers' online review data by employing both sentiment analysis methods for diagnosis and co-occurrence analysis for prescription.	-	Theoretical: A step-by-step guideline for future research by demonstrating a hybrid text mining process. Practical: Positive and negative aspects about Paris. Sentiment analysis of unsatisfied customers of Paris transportation.
Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach	Sustainability	Gang Ren, Taeho Hong	2017	China	Methodological	Topic based sentiment analysis	latent Dirichlet allocation (LDA)	Bazhuayu (website crawler)	Ctrip.com	n= 2,000 reviews	A Research framework that can make most of the outcomes (word-topic matrix, document-topic matrix) of the LDA model		Theoretical: Classifying sentiments in topics into emotions Practical: Advice for destination managers in general. Insights about sentiments and discussed topics of Chinese Tourists in China.

Appendix 6: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 6)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
Online destination image: Comparing national tourism organisation's and tourists' perspectives	Tourism Management	Athena H. N.Mak	2017	Taiwan	Tourism management	Visual Content Analysis. Content Analysis	Mise en scene analysis Perceptual mapping Analysis	Nvivo Software (for visual content analysis and mise en scene analysis). CATPAC II software (for textual content analysis). ThoughtView software (for 3D perceptual maps for the TGC and the NTO datasets)	travelpod.com travelblog.org	n= 1,953 photos and 140 pages of textual data	-	-	Theoretical: Travel photos are capable to depict cognitive DI attributes and to embody the affective DI attributes of places. Practical: DI of Eastern Taiwan. Some cognitive DI dimensions identified in the TGC data were found to be under-represented in the NTO data
Understanding the creation of destination images through a festival's Twitter conversation	International Journal of Event and Festival Management	Lluís Garay Soledad Morales Pérez	2017	Spain	Tourism management	Qualitative and quantitative Content Analysis	Cognitive affective conative classification. Speech-act-analysis.	* Nvivo software (mining the Twitter application programming interface (API)) * Ncapture (Nvivo add-on tool to capture Twitter data)	Twitter	n= 3,226 tweets and retweets t= 3,5 years	-	-	Theoretical: Twitter is a helpful source to find all kind of DI agents. Practical: Active conversations on Twitter can lead to more visitors. Speech and User analysis.
Understanding destination image from the perspective of Western travel bloggers: the case of Istanbul	International Journal of Culture, Tourism and Hospitality Research	Zaid Alrawadieh, Mithat Zeki Dincer, Fusun Istanbulu Dincer, Parvin Mammadova	2018	Turkey	Tourism management	Qualitative Content Analysis	-	-	TravelPod.com	n= 141 posts about Istanbul t= 1 year	-	-	Practical: DI of Istanbul by Western tourists. Concrete advice for DMOs.
Feeling a destination through the “right” photos: A machine learning model for DMOs’ photo selection	Tourism Management	Ning Deng, Xiang Robert Li	2018	USA	Tourism management	Content Analysis. Sentiment Analysis	Cognitive Affective Concept	Python-based Flickr API (to crawl the viewers’ comments) SentiWordNet (to measure the sentiment values of adjectives)	Flickr	n= 20,974 photos t= 10 years	-	-	Practical: Machine learning model for picking the right UG photos to shrink the gap between projected and perceived image, and find photos to illustrate emotions

Appendix 7: XYZ DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 7)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
Is Xenios Zeus Still Alive? Destination Image of Athens in the Years of Recession	Journal of Travel Research	Alkmini Gkritzali, Dimitris Gritzalis, Vassilis Stavrou	2018	Greece	Tourism management	Sentiment Analysis	semiautomated analysis web content mining	* SentiStrength software (for sentiment analysis) * Weka software and Apache Mahout (for sentiment rating and text mining)	TripAdvisor.com	n= 109,460 messages from the Athens Travel Forum t= 10 years	-	safe environment, hospitable environment, general mood and atmosphere, relaxing effect, authenticity of experience	Practical: Perception of Athens varies between Greek and non-Greek and between time periods. Crisis should be seen as potential new beginnings in tourism
Finding patterns in urban tourist behaviour: a social network analysis approach based on TripAdvisor reviews	Information Technology & Tourism	Egbert Van der Zee, Dario Bertocchi	2018	Belgium	Tourism management	Network analysis	Content pattern visualization	* Webbased scraping software Kimono (for finding reviews on Antwerp and match them to user profiles) * UCINET (for conducting the social network analysis and visualization)	TripAdvisor.com	n= 352,790 reviews	-	-	Practical: Tourism movement around Antwerp. Geographic clustering should be considered more in tourism marketing.
Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis	Tourism Management	Estela Marine-Roig, Berta Ferrer-Rosell	2018	Spain	Tourism management	Quantitative Content Analysis	Compositional analysis. Spatial, cognitive, and affective components	-	TripAdvisor.com. Press kit from the Catalan Tourist Board. Lonely Planet guidebook on Catalonia.	n= 80,000 reviews	-	Methodology for assessing TDI gaps between projected and perceived images in relation to the destination image components.	Theoretical: Perceived DI differs strongly from projected DI Practical: Perceived and projected image of Catalonia and Barcelona
Consumer Perception of Tourist Experience through Online Reviews: The Islands of the Senses of Cape Verde	Worldwide Hospitality and Tourism Themes	Oliveira, C., Brochado, A., Moro, S. & Rita, P.	2019	Cape Verde	Tourism management	Automated Content Analysis	-	Leximancer software (text mining)	TripAdvisor.com	n= 472 reviews	-	-	Practical: Most common themes in online reviews about Cape Verde Islands, with detailed description.

Appendix 8: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 8)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
Are Social Media Data Pushing Overtourism? The Case of Barcelona and Chinese Tourists	Sustainability	María-del-Mar Alonso-Almeida, Fernando Borrajo-Millán, Liu Yi	2019	Spain	Tourism management	Sentiment Analysis	Co-occurrence analysis. Network-analysis	LIWC (Linguistic Inquiry and word count) Gephi (clustering analysis program) Web crawler tool	Chinese online travel agencies websites: Baidu Travel, Qunar, Mafengwo, Ctrip	n= 11,655 tour reviews of Barcelona city	-	-	Practical: positive and negative aspects about Barcelona from a Chinese perspective. Behavior of Chinese tourists in Barcelona.
Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations	Tourism Management	Yi Liu, Kaixuan Huang, Jigang Bao, Kaiqi Chen	2019	Australia	Tourism management	Sentiment Analysis	Lexicon filtering. Co-occurrence analysis.	LIWC (Linguistic Inquiry and word count) Gephi (clustering analysis program) Web crawler tool	TripAdvisor.com, Qyer.com, Mafengwo.com, Ctrip.com	n= 36,148 online travel reviews t= 3 years	-	-	Practical: Preferences of Chinese tourists differ highly from international tourists. Chinese tend to be more critical. Chinese tourists' positive sentiments were associated with the richness of tourism resources and the urban function of the destinations.
Is hosting tourism events a sustainable way of destination branding: Evidences from Sanya, China	Cogent Social Sciences	Yibin Chen, Yi Liu, Chen Wang, Kaiqi Chen, Willem Coetzee	2019	China	Tourism management	Quantitative Content Analysis	Tourists' sentiment evaluation model (TSE)	* Python programming (to calculate the score of positive and negative tourism emotions) * Webcrawler software	TripAdvisor.com, Baidu Travel, Mafengwo, Qunar, Ctrip	n= 129,791 reviews t= 3 years	-	-	Practical: Tourists tend to write more positive reviews (Pollyanna effect). Sport events tend to be effective to improve the DI.
Destination Image Analytics Through Traveller-Generated Content	sustainability	Estela Marine-Roig	2019	Greece	Tourism management	Quantitative Content Analysis	Sentiment analysis. Attitudinal and behavioral response.	Algorithm created by authors	TripAdvisor.com	n= 300,000 reviews	-	-	Theoretical: An initial integrated framework for analytics on TDIs from a massive amount of TGC. Practical: Top 12 most frequent words about Attica in UGC. Most popular attractions/ restaurants / hotels. Concrete advises for policy and destination managers in Attica.

Appendix 9: DATABASE DEVELOPED AND JOURNALS INCLUDED. SHORT VERSION (PART 9)

Title	Journal	Researchers	Year	Country	Approach	Methods used			Analyzed Media	Sample Size	Results		
						Main method	Additional methods / applied theories	Technology /Programs			models developed	categories used / found	theoretical or practical advice
A Psycholinguistic View of Tourists' Emotional Experiences	Journal of Travel Research	Kamal Rahmani, Juergen Gnoth, and Damien Mather	2019	New Zealand France Germany Italy Brazil, Mexico Iran Japan Thailand Egypt	Linguistic	Quantitative Content Analysis	Structural Equation Modeling. Semantic Analysis	SAS text-miner add-on	TravelBlog.org	n= 1,000 travelblog entries	Model on Tourists emotional experience	-	Theoretical: Methodology to measure experience
#Visitspain. Breaking down affective and cognitive attributes in the social media construction of the tourist destination image	Tourism Management Perspectives	Lluís Garay	2019	Spain	Marketing	Quantitative Content Analysis	affective and cognitive attributes. Product and place dimensions	Ncapture® (software that captures and structures tweet data)	Twitter	n= 1,455 tweets t= 1 year	-	-	theoretical: usefulness to profile stakeholders, break down affective and cognitive attributes and consider seasonality in destination image
Symmetry Analysis in Analyzing Cognitive and Emotional Attitudes for Tourism Consumers by Applying Artificial Intelligence Python Technology	Symmetry	Guoxia Sun	2020	China	Tourism management	Quantitative Content Analysis	Co-occurrence analysis. Cognitive & Emotional Analysis	Python	Ctrip.com mafengwo.cn tuniu.com	n= 219,426 words of travel journals and 28,441 words of reviews	-	-	Practical: DI of Jilin province and precise suggestions for the DMO of Jilin.
TDIVis: visual analysis of tourism destination images	Frontiers of Information Technology & Electronic Engineering	Meng-qi CAO, Jing LIANG, Ming-zhao LI, Zheng-hao ZHOU, Min ZHU	2020	China	Tourism management	Sentiment Analysis	Keyword-based sentiment visualization method. Latent Dirichlet allocation (LDA)	TDIV program (Tourist Destination Image Visualization)	Mafengwo.com Baidu Travel	n.a.	Proposed a system to help users build a relatively complete destination image with the help of TDIV	-	Practical: Visualized DI of Chengdu City.