



Web Reputation of Protected Areas as Tourism Destinations: a Case Study

Annalisa Stacchini¹

Abstract

The purpose of this study is to investigate the reasons why tourist inflows in three Italian protected areas, endowed with valuable natural and cultural heritage, are below their potential. These sites participate in the European Regional Development Project EXCOVER, that aims at developing sustainable tourism in underrated Adriatic areas. To this goal, it is important to investigate the reputation of the sites among the general public, especially in the imaginary of people who have never visited it, to understand and change what keeps tourists away. This task is brought about through a semantic analysis of online reviews, that are crucial in the construction of non-visitors' imaginary about a destination and influence travel choices. Customers' reviews are scraped from TripAdvisor through a code written on-purpose. Topics are modelled through an unsupervised machine learning algorithm, chosen because a consolidated theoretical model of destination reputation is still missing, so reviews are allowed to tell what is relevant to reviewers, without prior constraints. Moreover, to the exploratory aim of this study, online opinions are not considered as answers to predetermined questions, but as free instances of word-of-mouth. The resulting proxies of web reputation are compared across time and space, to get the information needed to devise effective tourism development strategies and marketing initiatives. The originality of this contribution lays in the investigation of the web reputation of protected areas still to be developed into tourism destination. Moreover, it may provide some hint useful to the formulation of a solid theoretical model of destination reputation.

Keywords: Destination reputation Semantic analysis Web reputation, Topic modelling Protected areas Sustainable tourism development



Annalisa Stacchini.) annalisa.stacchini2@unibo.it

Cite this article

Stacchini A. (2020) Web Reputation of Protected Areas as Tourism Destinations: a Case Study. *EATSJ - Euro-Asia Tourism Studies Journal*, Vol.1, Issue: December 2020 pp. 106-120.

Received: 20 July 2020 | Accepted: 12 November 2020 | Published online: 16 December 2020

¹ University of Bologna, Italy

Introduction

The intangible nature and the value-creation capability of a robust reputation make it a hardly imitable key source of competitive advantage (Quintana-García, Benavides-Chicón & Marchante-Lara, 2020). With reference to tourism destinations, it has been shown that a strong reputation improves tourists' attitudes, perceived value and loyalty, destination's image, tourist satisfaction and intention to visit (Marchiori, Cantoni and Fesenmaier's, 2013). Whence the importance of reputation analysis, especially for areas still to be developed into tourism destinations: by detecting its determinants and how they are perceived, it can lead destination managers' decision-making processes to the improvement of the local tourism supply and, consequently, of the destination's reputation and competitiveness. In fact, it has been shown that a positive reputation is built by taking good actions and optimizing the locally available resources and capabilities, rather than through marketing and communication campaigns (Burke, 2011).

Most extant literature investigates destination's reputation with tourists (e.g. Artigas, Vilches-Montero & Yrigoyen, 2015; Widjaja, Khalifa & Abuelhassan, 2019), based on the hypothesis that familiarity with a place is necessary to form a personal perception of its reputation. But people who have never visited a destination use multiple second-hand sources, functioning as reputation mediators, to get a sense of the dominant public opinion, on which to base their travel choices (Dickinger, 2011). Nowadays, online word of mouth (WoM) has become transparent, as users share their opinions and experiences on social media, travel blogs and primarily on Online Travel Agencies (OTAs), where they can read customers' reviews, but also book tourism services and find information published by the supply side. The analysis of user-generated contents is a feasible way to investigate destinations' reputation with the general public, as online reviews can be considered proxies of the dominant public opinion and have been claimed to influence the intention to visit (Dickinger, 2011).

Thus, in this study a web reputation analysis is carried out to detect the reasons why tourist inflows in 3 Italian protected areas, endowed with valuable natural and cultural heritage, are well below their potential. These sites participate in the European Regional Development Project EXCOVER, that aims at developing sustainable tourism in underrated Adriatic areas. To this goal, it is crucial to investigate the reputation with the general public, to understand and change what keeps travellers away. This task is brought about through a completely probabilistic Bayesian semantic analysis of 5,462 reviews published on TripAdvisor. A Latent Dirichlet Allocation model is used to detect the salient topics describing web reputation, because it yields easily interpretable and efficient representations topics, also for small corpora (Deepak, 2017). Being completely data-driven, this method allows to identify strengths and weaknesses peculiar of each destination's reputation, without limiting the analysis to themes considered important by the researcher (but possibly not by the general public), as it happens in sample surveys and partially probabilistic topic modelling algorithms (Marchiori et al, 2013).

This paper continues with a review of the main relevant literature. The empirical setting and the analytical method are described in section 3. Afterwards, results and implications are discussed, followed by concluding remarks.

Literature Review

Reputation analysis was introduced in corporate studies to understand observers and stakeholders' collective judgment of a corporation (Barnett & Hoffman, 2008). Based on converging definitions, (e.g. Fombrun, 1996; Wartick, 1992), since the eighties the corporate literature developed a solid theoretical framework and was enriched by many empirical contributions, confirming the importance of this construct (e.g. Fombrun & Shanley, 1990; Gray & Balmer, 1998). Conversely, although reputation is even more crucial for destinations than for single companies (Marchiori et al, 2013), to date just few and quite divergent attempts have been made to set a theoretical model of destination's reputation. Mingchuan (2015) defines the latter as a collective "general feeling, impression and cognition" (p. 35) shared among the public, based

on past experiences. Artigas et al (2015) underline the cumulative and dynamic nature of this construct. Morgan (2011) articulates reputation in the dimensions of conversation, discrimination and differentiation. Prado and Trad (2012) identify reputation with the degree of consistency of destination's image and identity. Inversini et al (2010) developed a destination reputation model, adjusting the RepTrak corporate framework, that was revised in Marchiori, Inversini, Cantoni and Dedekind (2019).

Most of the empirical studies on destination's reputation are based on tourist surveys (e.g. Su & Huang, 2012; Artigas et al, 2015; Widjaja et al, 2019), that allow to investigate the perceptions of tourists only, but may shed some light also on the reputation with the general public, because people who never visited the place form their perception of its reputation also based on tourists' opinions and stories. However, the general reputation is also founded on visitors, commuters and residents' WoM, which nowadays are increasingly published on the web 2.0, which enables consumers to provide, collect and rate information about experiences, products and services (Nicholas, Huntington, Jamali & Dobrowolski, 2007). "The aggregation of the entire range of online representations creates the web reputation of organizations" (Inversini et al, 2010, p. 323). Many studies prove that online users' reviews are becoming the most trusted source of information based on which travellers choose their next destination and related services (e.g. Dickinger, 2011; Manaman, Shahram & Abolfazl, 2016). Whence the importance of web reputation analyses, as the starting step for devising an effective marketing strategy.

Most web reputation studies rely on sentiment analysis (e.g. Lai & Ming, 2015; Vu, Li, Law & Zhang, 2019), that consists in modelling the polarity of user-generated contents (UGC). This analysis is sound especially for non-binary semantic traits, but does not help to understand non-polar, abstract and life experience-related concepts (Rahmani et al, 2019). Conversely, semantic analysis allows to wholly investigate the meaning of a text, by reconstructing the topics therein contained (Marchiori et al, 2013). Two main strategies for topic modelling can be employed. The partially probabilistic approach consists in computing functions of the frequencies of predefined topic lexica in the text (see: Stone, Dunphy & Smith, 1966). Conversely, the completely probabilistic topic modelling framework is data-driven, so text topics are not defined *a priori*, but detected through unsupervised machine learning algorithms, generally Latent Dirichlet Allocation Bayesian model (LDA, see: Blei 2012) or Singular Value Decomposition (SVD, see: Dumais 1991). Partially probabilistic methods treat UGC as answers to an implicit survey, so they are appropriate to test theoretical models, or to investigate 'general' phenomena, like emotions or purchasing behaviours (Marchiori et al, 2013). While completely probabilistic models allow a free exploration of text meanings.

Given the fundamental importance of web reputation for tourism destinations and its ease of access, it is surprising that very few empirical studies address this topic, that is nonetheless increasingly explored with reference to accommodations (Phillips, Barnes, Zigan & Schegg, 2017) and restaurants (e.g. Vu et al, 2019). Inversini et al (2010) examine the web reputation of London through a descriptive analysis of relevant UGC from 95 websites, returned by a Google search. A very similar analysis is carried out by Marchiori et al (2019). The present work adds a new contribution to the seminal empirical literature about destinations' web reputation, focusing on 3 Italian protected areas.

Methodology

Empirical setting

The areas under consideration are substantially unknown, both nationally and internationally. They all suffer from depopulation and lack of opportunities for young people. Alfonsine is a municipality of 106,79 square kilometers, with a population of 11,993 inhabitants in the hinterland of lowland Ravenna province. It hosts a natural reserve, included in the Po Delta Park, constituted by woods, valleys and marshes. Ostellato is a municipality of 173,34 square kilometers, with a population of 6,030 inhabitants, in the center of Ferrara province. Its natural heritage includes the Valleys, a natural reserve with plant and animals typical of wet freshwater environments, and the Mezzano, a vast reclaimed agricultural area, with large populations of sedentary fauna. Sasso Simone and Simoncello is an interregional park of 49.91 square kilometers in the heart of the historic Montefeltro region, characterized by isolated mountains and rocks, and a flourishing wildlife.

As the assessments of local products and services have been found to contribute the most to overall reputation in survey-based structural models (Widjaja et al, 2019), 5,462 users' reviews regarding local attractions, hotels and restaurants were scraped from TripAdvisor, using Python, Selenium package. Their distribution by object and destination is shown in table 1. The reviews were automatically translated using Google Translate.

Semantic analysis

As the aim of this analysis is discovering what the online WoM is about, with no predetermined focus, also because of the absence of a solid theoretical background, a completely probabilistic approach is adopted. Examining proxies of web reputation permits to detect the local peculiarities forming a destination's identity and making it unique in the imaginary of the general public. Thus, attractions, hotels and restaurants reviews are analyzed through a LDA model, because it yields easily interpretable and more efficient topics representations than SVD method, also for smaller corpora (Deepak, 2017). In the LDA Bayesian framework, topics are unobservable variables producing observable reviews. Given a collection of M documents (retrieved from TripAdvisor) $d_j, j=\{1, 2, \dots, M\}$, where N words $w_i, i=\{1, 2, \dots, N\}$ out of W occur, the Bayesian approach assumes that the same term can be used to express K different meanings of each topic $t_k, k=\{1, 2, \dots, K\}$, with probability $P[w_j|t_k]$. The most probable topic meaning is selected based on the posterior probability that the observed words express that meaning $P[t_k|w]$. The frequency of each word, $n_w=(n_{w_1}, n_{w_2}, \dots, n_{w_W})$ follows a Multinomial($N, P[w|t_k]$) distribution:

$$P[n_w|p] = \frac{N!}{n_{w_1}!n_{w_2}!\dots n_{w_W}!} \prod_{i=1}^W p_i^{n_{w_i}} \quad (1)$$

where $p=(p_1, p_2, \dots, p_D)=P[w|t_k]=(P[w_1|t_k], P[w_2|t_k], \dots, P[w_w|t_k])$. As the same word can be used to write K different topic meanings, also $p=P[w|t_k]$ takes different values with different probability $P[p_k]$. The natural conjugate (prior distribution of the parameter p , belonging to the same distribution family of the posterior) of the Multinomial distribution is the Dirichlet:

$$P[p] = \frac{1}{B(\alpha)} \prod_{k=1}^K p_k^{\alpha_k - 1} \quad (2)$$

where $P[p]=P[P[w|t_k]]$, can be interpreted as the continuous-case equivalent of (number of all the possible simple combinations of scraped words' frequencies), as it is the beta function applied to α vector of hyperparameters, representing the prior frequency of each topic meaning (see: Griffiths & Steyvers, 2004). By the Bayes' theorem, the posterior distribution $P[p|w]=P[P[w|t_k]|w]$ is a Dirichlet():

$$P[p|w] = \frac{1}{B(\alpha+n_w)} \prod_{k=1}^K p_k^{\alpha_k+n_w-1} \quad (3)$$

Then, the topic meaning most probably conveyed by the terms analysed is that for which $P[p|w]$ reaches the maximum value. Equations 1-3 show LDA assuming that reviews deal with only one topic, but each word can express multiple topics, thus the probability functions must be added for each topic in a mixture:

$$P[P[w|t_1, t_2, \dots, t_T]|w] = P[p_{\{t\}}] = \sum_{t=1}^T n_{w,t} \frac{1}{B(\alpha+n_w)} \prod_{k=1}^K p_{t,k}^{\alpha_{t,k}+n_w-1} \quad (4)$$

where $n_{w,t}$ indicates that all T topics are considered and is the word frequency, that weights the t -th topic in the mixture (which can be read as a weighted average of posteriors). Similarly, each review is a collection of words: $d_j=(w_1, w_2, \dots, w_{w_j})$, $P[d|t_k] \sim \text{Multinomial}(M, P[d|t_k])$, $P[d|t_k] \sim \text{Dirichlet}(\alpha)$ and $P[P[d|t_k]|d] \sim \text{Dirichlet}(\alpha)$. A same review can contain different topics, so a mixture is built as for words. Finally, the posterior probability that, for every topic, $t_{t,k}$ is actually expressed by the words written in the documents scraped is:

$$P[P[d|t], P[w|t]|d, w] = P[P[w|t]|d, w, P[d|t]]P[P[d|t]|d, w] \quad (5)$$

The model is estimated with the Gibbs sampling algorithm (see: Griffiths & Steyvers, 2004). Maximum a posterior estimates (MAP) of are discrete probability values, each associated with a meaning, that can be described by the terms with the highest likelihood to be observed, given that topic, , and the most probable pair of words. LDA can be carried out with various software, in this work R packages tm and textmineR are used to clean the data, make count statistics and estimate the LDA model.

Results

Data description

Interesting preliminary insights can be drawn from the distribution of the scraped reviews by time, object and destination, consistently with the conception of reputation as a cumulative and dynamic construct (Artigas et al, 2015). Looking at table 1 and interpreting the number of reviews in terms of popularity (as in Viglia, Furlan & Ladrón-de-Guevara, 2014), Sasso Simone appears the most popular destination among these protected areas, thanks to its wide and high-quality restauration offer, that should be leveraged by local destination marketing managers as a main driver of tourism development (Vu et al, 2019). The most popular attractions are in Alfonsine, unexpectedly, as both residents and local experts expressed skepticism about the capability of the events and assets present in the area to attract tourists. Thus, this datum is of great relevance for decision makers involved in the tourism development of Alfonsine, as it points out that important attractions have been overlooked, while they might be well preserved and effectively promoted to strengthen the competitiveness of this small area as a tourism destination. Overall, the number of restaurants reviews is by far the highest, because they include also those by residents, commuters and travellers just passing by. Nonetheless, they can help developing the local tourist demand, as satisfying food & wine experiences may be able to communicate wider attractive characteristics of a destination, especially with reference to cultural impressions and insights (Andersson, Mossberg & Therkelsen, 2017).

Table 1. Number of reviews by object and destination. Source: data scraped from TripAdvisor, elaboration by the author.

	Attractions reviews	Hotels reviews	Restaurants reviews	Total
Alfonsine	159	156	928	1243
Ostellato	68	333	1427	1828
Sasso Simone	75	71	2245	2391
Total	302	560	4600	5462

Figure 1 shows the dynamics of the yearly published reviews (those written before 2009 and in 2020 are left out, because the formers are very few/none and for the latter a few months are available). Interestingly, although the 3 areas start from the same level, their popularity has evolved very differently since 2012, when two major earthquakes struck the region. The popularity of Alfonsine grew quickly, with a peak in 2016, when the local football team was promoted in D series, then declined quite sharply. The capability of sport events to boost positive WoM and loyalty (see: Girish & Lee, 2019) should be taken into account by local destination managers of all the developing destinations, as hosting an important contest could be a fast and very effective means to improve the area's visibility and attractiveness.

The 'fame curve' grew slower for Ostellato, reaching the top in 2017. As no particular event took place in that year, local policy makers should analyze possible modifications intervened in the

marketing and communication strategy, that could account for this loss of popularity, and correct it, as it appears to be currently worsening the visibility of this natural oasis. Conversely, the number of published reviews about Sasso Simone increased every year, at an especially fast pace between 2015 and 2017, after the regional administrative elections. This evidence suggests that policy changes, even just at the administrative level, can have a significant impact on the visibility of destinations (Kim & Kim, 2020).

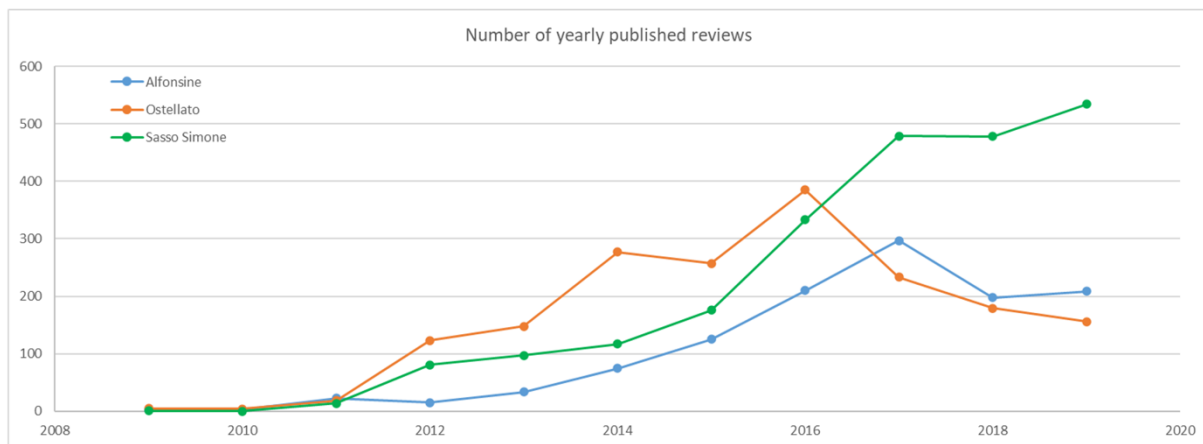


Figure 1 – Number of reviews by year of publication. Source: data scraped from TripAdvisor, elaboration by the author.

Although not being talked about is often worse than being talked badly (Klapp, 1964), this might not be true for destinations, because negative reviews may discourage more strongly travellers from choosing a place, than absence of reviews. Therefore, if the number of UGC can quantify the popularity of a destination, a text analysis is needed to assess its quality and detect its drivers (Viglia et al, 2014).

LDA estimation results

The LDA specification search was carried out comparing R-square values (measuring goodness-of-fit as in linear regression), adjusted for the number of topics, to determine the most meaningful UGC pooling level and the number of topics itself. A separate model is estimated for attractions, hotels and restaurants in each destination. Results are reported in tables 2-4. Along with the most probable terms expressing each topic, its coherence and prevalence is shown. Coherence is computed as the weighted average of the estimated proportion of times a word occurs together with another one of the same topic, over the single word's frequency. It measures the meaningfulness of the topics, the degree of usage-based semantic similarity between the most frequent terms. Prevalence is computed as the proportion of reviews dealing with a certain topic, over the total number of documents-topics. It measures the load of the topic on the web reputation. Models' R-squared are quite low, likely due to omitted variables and the small number of documents compared to that of different words employed.

Looking at the topics prevailing in the electronic WoM about Alfonsine, the most appreciated attraction is the labyrinth, built by the owner of a local agriturismo inside a corn field, that is the widest in Europe. This evidence confirms the importance of private businesses for sustainable tourism development (MacKenzie & Gannon, 2019). The second main driver of reputation is the kindness of the local businesses' owners and workers, as it might be expected in the light of some extant literature (e.g. Powell, Wang, O'Neill, Dentice & Neill, 2019). The cultural heritage looks important in framing the web reputation of this destination (unsurprisingly, see: Marchiori et al, 2013), where international tourists are happy to find guides speaking good English. Hotels are appreciated for the rooms' cleanliness, excellent service and great breakfasts, boosting intentions to return, in line with previous research (e.g. Phillips et al, 2017). Most reviews refer to 4 stars hotels: the standard in many Italian business travel policies (Guizzardi, Monti & Ranieri, 2016). In fact, most hotel clients are business travellers related to large factories in Alfonsine. The typical

enogastronomy, made with high quality ingredients, along with outstanding service, nice staff and fair prices make this place highly recommended. Traditional dishes play the starring role (as expected, see: Vuković & Terzić, 2020), especially cappellacci with pumpkin and fresh fish from the valleys. Among the top terms there is a single negative word: 'bad', likely referred to restaurants' interns.

Table 2. LDA estimation output for Alfonsine. Source: data scraped from TripAdvisor, elaboration by the author.

Alfonsine – Attractions							
topic	coherence	Prevalence	top_terms				
evening_friends	0.148	15.86	evening	labyrinth	Barbecue	dinner	friends
corn_field	0.116	13.54	labyrinth	year	Hours	idea	wonderful
lot_fun	0.144	13.17	children	nice	Beautiful	labyrinth	Fun
Mr_Galassi	0.059	11.35	kind	staff	Good	dinner	excellent
ephemeral_labyrinth	0.084	10.27	beautiful	labyrinth	experience	fun	Nice
entrance_labyrinth	0.095	10.19	labyrinth	entrance	Find	pizza	hours
wonderful_place	0.049	9.40	place	bring	Good	wonderful	Time
piece_history	0.170	7.91	church	inside	Sanctuary	center	world
world_war	0.279	5.87	museum	structure	Visit	room	history
great_place	0.233	2.44	museum	area	Visit	fantastic	english
Adj. R2	0.140						
Alfonsine – Hotels							
topic	coherence	Prevalence	top_terms				
clean_rooms	0.023	24.28	rooms	breakfast	Excellent	clean	staff
stayed_night	0.012	16.36	excellent	beautiful	Service	owner	return
4star_hotel	0.101	12.99	rooms	4stars	Parking	good	structure
hotel_located	0.062	11.91	breakfast	rooms	Large	free	town
4star_hotel	0.075	11.90	rooms	night	Breakfast	bathroom	Bed
4star_hotel	0.066	11.71	time	stay	Rooms	make	work
Adj. R2	0.112						
Alfonsine – Restaurants							
topic	coherence	Prevalence	top_terms				
excellent_quality	0.040	13.67	piadina	excellent	Quality	dishes	ingredients
highly_recommended	0.032	13.21	excellent	place	Nice	staff	dishes
excellent_service	0.016	12.17	excellent	service	Good	fish	large
good_quality	0.045	11.37	good	quality	Dishes	price	lunch
place_nice	0.057	11.13	dishes	place	Table	bad	Time
fresh_fish	0.036	9.81	fish	excellent	Good	fresh	dinner
pizza_good	0.040	9.77	pizza	good	Place	beer	Nice
pumpkin_cappellacci	0.081	8.77	pumpkin	excellent	cappellacci	typical	good
cold_cuts	0.131	7.80	good	cold	Dinner	tigelle	Cuts
fair_prices	0.254	2.30	prices	fair	Bottles	wines	Real
Adj. R2	0.088						

While the natural reserve in Alfonsine is not mentioned among the keystones of its web reputation, those in Ostellato are definitely pivotal. The Valleys and the Mezzano are described as peaceful and beautiful natural oases, full of animals, unfortunately including ubiquitous mosquitos. Reviewers enjoyed also the Valleys' swimming pool, an example of how private businesses can help valorizing natural reserves as tourist attractions (MacKenzie & Gannon, 2019). The temporary exhibition of life-size dinosaurs, hosted by the Museum of the Territory in 2014, left an important positive mark on the online reputation of Ostellato. As the Museum of the Territory is normally neglected, this evidence confirms that thematic events, can spark interest on

otherwise unattractive cultural assets (Gilmore & Rentschler, 2002). Excellence of hotels' personnel, location in the beautiful city center and breakfast characterize the electronic WoM about accommodations. The importance of these elements has already been highlighted by extant studies (e.g. Phillips et al, 2017). The typical fish, cooked in different ways, wide choice of piadina, pizza and ice cream stand out, besides fast service and nice staff. Business travellers seem to enjoy lunch at a good price there, confirming the price's impact on business consumers' experience found by previous literature (Faizan, 2015).

Table 3. LDA estimation output for Ostellato. Source: data scraped from TripAdvisor, elaboration by the author.

Ostellato – Attractions							
topic	coherence	prevalence	top_terms				
beautiful_place	0.091	23.52	place	nature	beautiful	fishing	Oasis
swimming_pool	0.376	13.45	pool	swimming	Swim	excellent	Nice
love_nature	0.149	13.29	nature	delta	Reeds	love	mosquitoes
swimming_pool	0.129	11.27	walk	nature	Time	nice	Long
surrounded_nature	0.083	10.35	oasis	peaceful	Animals	surrounded	Beautiful
beautiful_place	0.105	9.84	park	place	beautiful	valleys	Entrance
swimming_pool	0.097	9.76	bike	oasis	Peace	path	Bar
beautiful_place	0.292	8.53	exhibition	dinosaurs	Children	life	Museum
Adj. R2	0.087						
Ostellato – Hotels							
topic	coherence	prevalence	top_terms				
excellent_breakfast	0.057	20.27	excellent	staff	Good	location	Breakfast
city_center	0.074	14.95	center	beautiful	City	rooms	Located
excellent_location	0.073	14.38	excellent	time	Place	family	Nice
excellent_breakfast	0.050	13.74	rooms	good	breakfast	clean	Stayed
hotel_located	0.038	12.10	rooms	stay	Place	staff	Breakfast
Adj. R2	0.418						
Ostellato – Restaurants							
topic	coherence	prevalence	top_terms				
excellent_quality	0.023	12.48	excellent	dishes	Staff	place	Pleasant
fast_service	0.028	12.36	excellent	fish	Good	service	Large
excellent_pizza	0.069	11.58	pizza	good	Nice	place	Staff
wide_choice	0.030	9.89	piadina	good	excellent	quality	Choice
grilled_fish	0.123	9.50	fish	grilled	Fried	mixed	appetizers
ice_cream	0.031	8.74	dishes	place	Home	ice cream	Great
business_lunch	0.040	8.26	lunch	menu	Price	good	Nice
raw_materials	0.529	4.03	materials	raw	ingredients	impeccable	Fair
Adj. R2	0.085						

Considering that reviews for Sasso Simone were searched by the name of the town where the Park Authority is based, extended to the suggested near places, it is worthy of note that the main topic shaping these small areas' reputation is actually this natural reserve. This result confirms the relevance of environmental sustainability to boost tourism development in a world that is increasingly more aware of the importance of preserving the environment and to favour green practices (Ashraf, Hou, Kim, Ahmad & Ashraf, 2020). However, the most recurring topic concerns the cultural heritage of the villages pointing the park, that tell the history of Montefeltro. This result is consistent with previous studies highlighting the centrality of cultural heritage to develop an attractive tourism supply in small green destinations (e.g. Dragouni & Fouseki, 2017). Finding an ancient printing house among the top terms describing Sasso Simone's web reputation confirms that local crafts are important elements of authenticity, highly valued by tourists (Nason, 1984). Excellent breakfast, welcoming staff's friendliness and clean rooms make

overnights stay near the park pleasant. As for Alfonsine, the only negative top term is 'bad', likely associated with rooms, considering that accommodation structures in Sasso Simone are quite old-fashioned. Restaurants appear highly recommended, thanks to excellent staff, attention to details, quality wines and ingredients, fair prices and impeccable service (elements known as drivers of restaurants' reputation, e.g. Vu et al, 2019). The wise balance of tradition and innovation, fundamental driver of product development in this industry (Petruzzelli & Savino, 2016), distinguishes the online reputation of the park area's restaurants.

Table 4. LDA estimation output for Sasso Simone and Simoncello. Source: data scraped from TripAdvisor, elaboration by the author.

Sasso Simone – Attractions							
Topic	coherence	prevalence	top_terms				
sasso_simone	0.131	18.75	visit	palace	building	Montefeltro	History
parish_church	0.275	13.92	church	parish	romanesque	pleasant	Time
sasso_simone	0.113	12.04	woods	walk	place	paths	Beautiful
sasso_simone	0.201	11.86	park	paths	path	nature	Villages
printing_house	0.354	11.75	printing	ancient	house	traditional	Made
sasso_simone	0.085	10.90	beautiful	good	years	great	Place
sasso_simone	0.426	10.47	beautiful	trail	path	rocks	Climb
sasso_simone	0.071	10.32	structure	wonderful	due	town	Places
Adj. R2	0.292						

Sasso Simone – Hotels							
Topic	coherence	prevalence	top_terms				
	0.056	17.34	staff	clean	excellent	breakfast	Friendly
excellent_breakfast	0.077	14.75	good	excellent	night	location	Small
clean_room	0.156	13.62	room	bathroom	breakfast	water	Shower
excellent_hotel	0.121	10.62	splendid	located	city	home	Welcoming
sasso_simone	0.088	9.70	place	pleasant	nice	bad	Rooms
Adj. R2	0.100						

Sasso Simone – Restaurants							
Topic	coherence	prevalence	top_terms				
highly_recommended	0.044	15.92	excellent	staff	dishes	recommended	Nice
ice_cream	0.080	12.25	ice	flavors	cream	shop	Excellent
attention_detail	0.046	10.64	dishes	time	quality	great	Place
nice_place	0.035	10.20	good	nice	price	reviews	Dinner
wine_list	0.083	9.67	dishes	evening	service	dinner	Pleasant
tradition_innovation	0.092	9.54	excellent	fish	menu	dishes	Innovation
raw_materials	0.120	9.36	excellent	pizza	raw	quality	Materials
fair_prices	0.364	3.04	fair	prices	impeccable	service	Good
Adj. R2	0.078						

Overall, these results are unexpected: the aim of this analysis was to discover reputational weaknesses preventing tourism demand to reach the levels expected given the quality of the local cultural and natural heritage. While the latter is very well represented in the reviewers' words, no serious deficiency was detected. Thus, regarding reputation, the reason of low tourist inflows seems the little popularity and visibility of these areas, which could be increased through effective web marketing actions (Pan, Xiang, Law & Fesenmaier, 2011).

Conclusion

This paper presented a semantic analysis of 5,462 reviews, published on TripAdvisor, regarding tourism-related attractions and services in 3 Italian protected areas, to investigate why tourist inflows are below their potentials, given the local endowment of valuable cultural and natural resources. These sites participate in the European Regional Development Project EXCOVER, aiming at developing sustainable tourism in underrated Adriatic areas. To this goal, it is important to investigate web reputation, because it shed some light on the imaginary of people who did not visit the destination (likely formed by reading online reviews), to understand and change what keeps tourists away. To the best of my knowledge, this is the first web reputation analysis of protected areas as tourism destinations. The scientific contribution brought by this work consists in showing how UGC can be employed as proxies of destination reputation among the general public, through the completely probabilistic Latent Dirichlet Allocation model. The method employed is inexpensive, more comprehensive and less researcher-influenced than sample surveys. It can be applied to investigate the web reputation of any destination and is especially useful in tourism development planning.

Results show that typical Italian food, quality service, friendly business owners and personnel are the pillars of the very positive reputations of these areas. Ostellato owes its strong reputation to natural reserves and seems to be effectively managing the tourism versus territory conflict (Almeida et al., 2018), balancing the environment protection with the business of a leisure structure, including a swimming pool and restauration services. Conversely, big factories in Alfonsine jeopardize the beauty and integrity of its natural oasis. In fact, many reviewers deal with business trips and the natural reserve does not appear among the most popular topics. The strength of this destination is the traditional enogastronomy. High quality ingredients make food experiences in all the considered areas satisfying and highly recommendable. According to the local farmers, the secret of their product is the unpolluted environment where crops and animals grow. Thus, protected areas can provide tourists with more than natural attractions, because a well-preserved environment yields healthy top-quality food (Lu, Song, Wang, Liu & Meng, 2015). The natural reserve of Sasso Simone and Simoncello is the main driver of the web reputation of the area, along with the huge cultural heritage disseminated throughout the territory. There, the protection of both the natural environment and the ancient historical heritage seems well-balanced with the capability to deliver valuable tourist experiences. The only negative aspects of the web reputations analysed refer to the interiors of hotels in Sasso Simone and restaurants in Alfonsine, possibly due to the low return on investments in the local tourism businesses, highlighted by some owners during Project meetings. Based on these results, the reason of low tourist inflows may be the little popularity and visibility of these areas, which could be risen through effective web marketing and communication actions (Pan et al, 2011), that are the acme of the Project.

These findings have relevant implications for regional policy-making oriented to sustainable development. In Alfonsine, local authorities should adopt a collaborative and facilitating approach towards entrepreneurs, who are now struggling between excessive bureaucracy and taxes, because the success of private businesses can boost the regional growth benefitting all the residents. By lowering the cost of investments, that is currently preventing business owner from improving their offer, more pleasant interiors and ambiances could be built, removing the only reputational weakness of this town. Moreover, the cultural heritage of Alfonsine, that is currently neglected as a driver of development, should be protected with care and valorized, as it has the potential to increase both the visibility and attractiveness of this area. More important, from the present analysis the protected area that theoretically should constitute the main driver of sustainable development, to grasp the opportunity represented by the growth of green tourism, appears completely unknown to online reviewers. Therefore, to date the endowment of this natural reserve looks just a cost and a constraint for the community of Alfonsine, that is drawing no benefit from it, in terms of tourist attractiveness and economic growth. Whence the need to devise marketing and communication actions aimed at making Alfonsine's natural reserve popular and charming, in the imaginary of the general public. Furthermore, outdoor sport events

and nature-related manifestations could be organized to help boosting the green reputation of this small area.

In Ostellato, policy makers should leverage the possibility to develop school tourism, thanks to the educational value of the great biodiversity for which the local natural reserves are mainly popular. To this aim, it would be fundamental to restore the Museum of the Territory, where to add interactive installations, organize educational workshops and events for schools and families. Once organically structured in a tourism offer reflecting the place identity, such attractions should be locally advertised, so that the many business travellers stopping in Ostellato for a good traditional meal would be aware of these occasions offered by the territory, beyond food & wine, and be induced to return with their family for a leisure trip.

The main suggestion for policy makers of Sasso Simone e Simoncello, deriving from this reputation analysis, is to keep going in this direction, with reference to the 'heritagization' of the territory (Dragouni & Fouseki, 2017) that is sustaining a strong reputation of the area as a green and cultural tourism destination. The focus on locally produced food & wine, characterizing the communication and marketing actions of Sasso Simone's destination managers looks successful and has the potential to drive the local trade, besides sustainable tourism. However, also here, as in Alfonsine, local businesses owners should be helped renovating their activities, to make the ambience more attractive and satisfying. To this aim, reducing taxes and bureaucracy could incentivize private investments, but also communicating the growing market opportunities that could be exploited by improving the accommodations appearance may be important, as many local entrepreneurs seem unaware of the high potential of this area and distrust the economic return they could draw from new investments.

The limitations of this study consist in the small number of UGC scraped (though all the available) and in the poor fit of some estimated models, that may be overcome by modelling space relations and time dynamics of reviews, or by adding covariates to the model, for example the reviewer's origin, numeric rating or trip type. For future works about destinations' web reputation it is recommended to investigate whether the elements emerged in this paper as fundamental components of this construct (food, natural and cultural heritage – but also private business-based attractions – service, kindness, prices, setting) could constitute a more general and valid measurement model, possibly helpful in the formulation of a solid theoretical framework of destinations' reputation.

Acknowledgments

I thank the ERDF for funding (grant Rep.10 Prot.n.52); EXCOVER Project Partners, destination managers and policy-makers for the useful information provided; the Center for Advanced Studies on Tourism of the University of Bologna, especially my supervisor, Andrea Guizzardi.

References

- Almeida J., Costa C., Nunes da Silva F. (2018). Collaborative approach for tourism conflict management: A Portuguese case study. *Land Use Policy*, 75, 166-179.
- Andersson, T.D., Mossberg, L. & Therkelsen, A. (2017). Food and tourism synergies: perspectives on consumption, production and destination development. *Scandinavian Journal of Hospitality and Tourism*, 17(1), 1-8.
- Artigas, E.M., Vilches-Montero, S. & Yrigoyen, C.C. (2015). Antecedents of tourism destination reputation: the mediating role of familiarity. *Journal of Retailing and Consumer Services*, 26, 147-152.
- Ashraf, M. S., Hou, F., Kim, W. G., Ahmad, W. & Ashraf, R. U. (2020). Modeling tourists' visiting intentions toward ecofriendly destinations: Implications for sustainable tourism operators. *Business Strategy and the Environment*, 29(1), 54-71.
- Barnett, M. & Hoffman, A. (2008). Beyond Corporate Reputation: Managing Reputational Interdependence. *Corporate Reputation Review*, 11, 1-9.
- Burke, J. (2011). Corporate reputations: Development, maintenance, change and repair. In R.J. Burke, G. Martin & C.L. Cooper (Eds.). *Corporate reputation managing opportunities and threats* (pp. 3-44). Surrey: Gower Publishing Limited.
- Deepak, S. (2017). A Survey on Journey of Topic Modeling Techniques from SVD to Deep Learning. *International Journal of Modern Education and Computer Science*, 9, 50-62.
- Dickinger, A. (2011). The Trustworthiness of Online Channels for Experience- and Goal-Directed Search Tasks. *Journal of Travel Research*, 50(4), 378-391.
- Dragouni, M., & Fouseki, K. (2017). Drivers of community participation in heritage tourism planning: An empirical investigation. *Journal of Heritage Tourism*, doi: [10.1080/1743873X.2017.1310214](https://doi.org/10.1080/1743873X.2017.1310214)
- Dumais, S. (1991). Improving the Retrieval of Information from External Sources. *Behavior Research Methods, Instruments & Computers*, 23 (2), 229-36.
- Faizan, A. (2015). International business travellers' experience with luxury hotel restaurants: the impact of foodservice experience and customer satisfaction on dining frequency and expenditure. *International Journal of Hospitality and Event Management*, 1, 164-186.
- Fombrun, C. J. (1996). *Reputation: Realizing Value from the Corporate Image*. Harvard: Harvard Business School Press.
- Fombrun, C., & Shanley, M. (1990). What's in a Name? Reputation Building and Corporate Strategy. *The Academy of Management Journal*, 33(2), 233-258.
- Gilmore, A. & Rentschler, R. (2002). Changes in museum management: A custodial or marketing emphasis? *Journal of Management Development*, 21(10), 745-760.
- Girish, V.G. and Lee, C.-K. (2019). The relationships of brand experience, sports event image and loyalty: Case of Jeju International Ultramarathon Race. *International Journal of Sports Marketing and Sponsorship*, 20(4), 567-582.
- Gray, E. R. & Balmer, J. M. T. (1998). Managing Corporate Image and Corporate Reputation. *Long Range Planning*, 5, 695-702.
- Griffiths, T. L. & Steyvers, M. (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5228-5235.
- Guizzardi, A., Monti, A. & Ranieri, E. (2016). Rating hotel quality for corporate business travel departments. *International Journal of Contemporary Hospitality Management*, 28, 2842-2863.
- Inversini, A., Marchiori, E., Dedekind, C. & Cantoni, L. (2010). Applying a Conceptual Framework to Analyze Online Reputation of Tourism Destinations. *Information and Communication Technologies in Tourism*, 8, 321-332.

- Kim B, Kim J. (2020). The rise of a new tourism destination: How did Vladivostok become the closest Europe for Korean tourists? *Journal of Eurasian Studies*, 11(2), 117-132.
- Klapp, O. E. (1964). *Symbolic Leaders*. Chicago: Aldine.
- Lai, L. & Ming, T. V. (2015). Content analysis of social media: A grounded theory approach. *Journal of Electronic Commerce Research*, 16, 138-152.
- Lu, Y., Song, S., Wang, R., Liu, Z. & Meng, J. (2015). Impacts of soil and water pollution on food safety and health risks in China. *Environment International*, 77, 5-15.
- Manaman, H., Shahram, J. & Abolfazl, A. (2016). Online reputation measurement of companies based on user-generated content in online social networks. *Computers in Human Behavior*, 54, 94-100.
- Marchiori, E., Cantoni, L. & Fesenmaier, D.R. (2013). What did they say about us? Message Cues and Destination Reputation in Social Media. In L. Cantoni & Z. Xiang (Ed.). *Information and Communication Technologies in Tourism*. Heidelberg, Germany: Springer.
- Marchiori, E., Inversini, A., Cantoni, L. & Dedekind, C. (2019). Towards a Tourism Destination Reputation Model. A first step. *Proceedings of the 6th International Conference Thought Leaders in Brand Management*, 921-930.
- Mackenzie, N. & Gannon, M.J. (2019). Exploring the antecedents of sustainable tourism development. *International Journal of Contemporary Hospitality Management*, 31(6), 2411-2427.
- Mingchuan, Z. (2015). The Reputation Crisis at the Tourism Destination: Connotation, Types and Management Strategies. *International Journal of Liberal Arts and Social Science*, 3(6), 35-40.
- Morgan, N. (2011). Destination brands: Managing place reputation. In A. Pritchard & R. Pride (Eds.). *Place Branding and Public Diplomacy* (3rd ed., pp. 316-320). London: Palgrave Macmillan.
- Nason, J.D. (1984). Tourism, handcrafts, and ethnic identity in Micronesia. *Annals of Tourism Research*, 11 (3), 421-449.
- Nicholas, D., Huntington, P., Jamali, H.J. & Dobrowolski, T. (2007) Characterizing and evaluating information seeking behavior in digital environment: spotlight on the bouncer. *Information processing and Management*, 43(4), 1085-1102.
- Pan, B., Xiang, Z., Law, R. & Fesenmaier, D. R. (2011). The Dynamics of Search Engine Marketing for Tourist Destinations. *Journal of Travel Research*, 50(4), 365-377.
- Petruzzelli, A. & Savino, T. (2016). Reinterpreting Tradition to Innovate: The Case of Italian Haute Cuisine. *Industry and Innovation*, 22, 677-702.
- Phillips, P., S. Barnes, K. Zigan & R. Schegg. (2017). Understanding the Impact of Online Reviews on Hotel Performance: An Empirical Analysis. *Journal of Travel Research*, 56 (2), 235-49.
- Powell, L., Wang, P., O'Neill, L., Dentice, G. & Neill, L. (2019). Webs of significance: Articulating latent value structures in a rural cafe organized as a worker cooperative. *Hospitality & Society*, 9(2), 183-213.
- Prado, F. & Trad, N. (2012). *Managing the reputation of places*. Reputation Institute. Retrieved from: <http://www.reputationinstitute.com/thought-leadership/white-papers>.
- Quintana-García, C., Benavides-Chicón, C. & Marchante-Lara, Macarena. (2020). Does a green supply chain improve corporate reputation? Empirical evidence from European manufacturing sectors. *Industrial Marketing Management*. Retrieved from: <https://www.sciencedirect.com/science/article/pii/S0019850119305929>.
- Rahmani, K., Gnoth, J., & Mather, D. (2019). A Psycholinguistic View of Tourists' Emotional Experiences. *Journal of Travel Research*, 58(2), 192-206.

- Stone, P. J., Dunphy, D. C. & Smith, M. S. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge: MIT Press.
- Su, L. J. & Huang, F. C. (2012). Study on the relationships among destination social responsibility, destination reputation, destination identification and tourism loyalty. *Tourism Tribune*, 10. Retrieved from http://en.cnki.com.cn/Article_en/CJFDTotol-LYXK201210014.htm.
- Viglia, G., Furlan R. & Ladrón-de-Guevara A. (2014). Please, talk about it! When hotel popularity boosts preferences. *International Journal of Hospitality Management*, 42, 155-164.
- Vu, H. Q., Li, G., Law, R. & Zhang, Y. (2019). Exploring Tourist Dining Preferences Based on Restaurant Reviews. *Journal of Travel Research*, 58(1), 149-167.
- Vuković, A.J. and Terzić, A. (2020). Gastronomy and Regional Identity: Balkan versus National Cuisine. In, A. Peštek, M. Kukanja & S. Renko (Ed.). *Gastronomy for Tourism Development*. UK: Emerald Publishing Limited, pp. 1-25.
- Wartick, S. L. (1992). The relationship between intense media exposure and change in corporate reputation. *Business & Society*, 31, 33-49.
- Widjaja, Y. I., Khalifa, G. S. A. & Abuelhassan, A. E. (2019). The effect of destination reputation on the revisit intention to Halal tourism destination of Jakarta. *International Journal of Business, Economics and Law*, 20(5), 104-111.