ISCTE Business School
Instituto Universitário de Lisboa

SENTIMENT ANALYSIS IN HOSPITALITY USING TEXT MINING: THE CASE OF A PORTUGUESE ECO-HOTEL

Ana Catarina dos Santos Calheiros

Dissertation submitted as a partial requirement for the conferral of Master of Science in Business Administration

Supervisor:
Prof. Dr. Paulo Rita, Full Professor, Marketing Department
ISCTE – University Institute of Lisbon
Co-supervisor:
Dr. Sérgio Moro, Researcher, Business Research Unit
(ISCTE-IUL)/ALGORITMI Centre (Univ. Minho)

December 2015

ISCTE Business School
Instituto Universitário de Lisboa

SENTIMENT ANALYSIS IN HOSPITALITY USING TEXT MINING: THE CASE OF A PORTUGUESE ECO-HOTEL

Ana Catarina dos Santos Calheiros

“Do not go where the path may lead, go instead where there is no path and leave a trail”

Ralph Waldo Emerson

Acknowledgments

I want to thank to my supervisors, Paulo Rita and Sérgio Moro, which were fundamental for the development of this dissertation, focusing on the availability and commitment they have shown to this project. Thanks for the entire learning and above all for challenging me at maximum.

To Areias do Seixo, namely to Gonçalo Alves, for providing me the opportunity to elaborate a project dissertation case study illustrated by a respected and highly esteemed Portuguese Eco-Hotel.

To my parents, friends and boyfriend to always be there when needed.

To EY for support me on this journey.

Abstract

The rapid development of the Internet and mobile devices enabled the emergence of travel and hospitality review sites, leading to a large number of customer opinion posts. While such comments may influence future demand of the targeted hotels, they can also be used by hotel managers for improving customer experience. Nevertheless, this trend poses a problem, considering information is widely scattered, making almost impossible to extract from it useful knowledge.

In this study, with the aim of facilitating this process, sentiment classification of an eco-hotel is assessed through a text mining approach using several different sources of customer reviews. Two dictionaries are compiled for building the lexicon used to parse the 401 reviews collected from a Portuguese eco-hotel between January and August of 2015. Then, the latent Dirichlet allocation (LDA) modeling algorithm is applied to gather relevant topics that characterize a given hospitality issue by a sentiment.

Findings of this study state that accuracy is influenced by interaction between LDA generated topic models and the correct construction of both dictionaries. These results also reveal that text mining can generate new insights into variables that have been extensively studied in hospitality industry, including that hotel food generates ordinary positive sentiments for the case studied, while hospitality generates both ordinary and strong positive feelings. Such results are valuable for hospitality management, validating the approach proposed.

Keywords: *Hospitality management; sentiment classification; text mining; customer reviews*

Jel Classification System:

Z32 Tourism and Development; **M30** Marketing and Advertising

Sumário

O rápido desenvolvimento da Internet e dos dispositivos móveis possibilitou o aparecimento de sites de viagens e sites de opinião na indústria hoteleira, levando a um grande número opiniões publicadas por parte do cliente. Embora, esses comentários possam influenciar a procura futura de certos hotéis, estes também podem ser usados pelos gestores dos hotéis para melhorar a experiência do cliente. No entanto, esta tendência representa um problema, uma vez que hoje em dia a informação se apresenta bastante ampla e dispersa, tornando quase impossível analisar todas as opiniões de clientes.

Neste estudo, com o objetivo de facilitar este processo, a classificação de sentimentos de um hotel ecológico é avaliada através de uma abordagem de "text mining" usando diversas fontes de comentários de clientes. Dois dicionários foram compilados para a construção do léxico usado para analisar os 401 comentários recolhidos a partir de um Eco hotel português entre janeiro e agosto de 2015. Em seguida, o algoritmo de modelação "latent Dirichlet allocation" (LDA) é aplicado para reunir tópicos relevantes que caracterizam uma determinada questão de hospitalidade por um sentimento.

Os resultados apurados neste estudo focam essencialmente que a precisão do mesmo é influenciada pela interação entre o modelo LDA, neste caso entre os tópicos por ele gerados e a correta construção de ambos os dicionários. Estes resultados revelam também que o "text mining" pode gerar novas perspetivas acerca de variáveis que têm sido extensivamente estudadas na indústria hoteleira, incluindo, no caso estudado, que a comida do hotel gera sentimentos positivos comuns, enquanto a hospitalidade gera ambos os sentimentos: positivos comuns e positivos fortes. Tais resultados são valiosos para a gestão hoteleira validando a abordagem proposta.

Palavras-chave: *Gestão hoteleira; classificação de sentimentos; text mining; opiniões de clientes*

Classificação JEL:

Z32 Tourism and Development; **M30** Marketing and Advertising

Abreviattions

BI = Business Intelligence

CRM = Customer Relationship Management

DM = Data mining

DSS = Decision Support Systems

e-WOM = Electronic word of mouth

LDA = Latent Dirichlet allocation

NLP = Natural Language Processing

SA = Sentiment analysis

SVM = Support vector machine

TM = Text mining

Table of Contents

List of tables

List of figures

1. Introduction

Hospitality traditionally lags other sectors in adopting information technology, but this has been changing in recent years and research into its multiple applications has followed suit (Šerić *et al*., 2014). With the unparalleled growth of Internet applications to travel and tourism, new information systems arise, producing a surplus of new opportunities and challenges to consumers and practitioners, making it an important sector and interesting field for research in different domains, such as management science, marketing and information technologies.

For hotels to survive in today's changing business environment, hotels' managers need to have a continuous focus on solving challenging problems and exploring new opportunities. That demands an investment in computerized support managerial decision making, which implies the need of decision support and business intelligence systems (Turban *et al*., 2011).

There is a lack of recent studies regarding BI applications into the main subjects of hospitality industry, thus motivating the present research. Although several studies have recently investigated text mining and sentiment analysis (e.g., Mostafa, M.M, 2013; Wang *et al*,. 2014), no previous studies have focused solely on investigating customer's sentiments towards a single hotel unit using text mining and LDA model. This study aims to fill this void.

Areias do Seixo was chosen as an example to illustrate that in a small hotel, with a different concept, as it is known as a "thematic luxury Eco-Hotel", it is also possible to apply BI systems to support hoteliers managerial decisions and enhance hotel effectiveness, and not only in major companies or hotel chains as it has been done already. The main goal is to transmit the availability of these BI applications across different sectors and the easy information access this DSS could provide in a near future.

This thesis presents an automated text mining analysis of customer reviews, regarding Areias do Seixo hotel, from January to August 2015 period, considering the sentiments expressed by customers towards hospitality issues. Moreover, the latent Dirichlet allocation topic modeling was used to cross both semantics of sentiment polarity and hotel domain in order to discover feelings generated by several hotel issues. This allows the identification of current customer trends and interesting future applications in this area. Furthermore, by focusing in several data sources, on-line and off-line texts, and by using both qualitative and quantitative methods to analyze the reviews, this study also adds breadth to the debate over hotel's quality as perceived by customers.

Lastly, findings of this study are expected to help both potential customers and the hotel industry to extract valuable knowledge from these reviews efficiently, in order to convert valuable data into actionable competitive intelligence and customer intelligence databases. Likewise, it can lead to a new trend of information processing in the tourism/hospitality industry, influencing the design of information systems in different functional areas.

More specifically, this research attempts to answer the following research questions:

RQ1: Can text mining techniques be used successfully to detect hidden patterns in consumer's sentiments in a specific hotel unit?

RQ2: Can sentiment mining techniques unveil how guest's satisfaction is being perceived, hence providing valuable knowledge for hotel managers' to understand the strengths and weaknesses of a specific unit?

This article is organized as follow. Next section provides a brief literature review on the major studies on text mining along with SA applications, also focusing on research concerning hospitality industry main trends and main BI applications. Section 3 describes the dataset used and also the method conducted for the analysis. In this section issues related to research, sampling and data analysis techniques are revealed. In section 4, the results are presented and discussed. Finally section 5 presents the implications and limitations of this study, and as give directions for future research.

2. Literature Review

"A literature review of a set of articles enables to analyze a given subject and identify trends of research and possible gaps that can lead to new studies and discoveries" (Levy & Ellis, 2006).

2.1 Hospitality industry

The Hospitality industry is crucial for the economy and thus it is a subject of great interest for researchers in different domains, such as management science, marketing and information technologies. Competition had also an effect on client related areas, with hotels increasing investment in customer retention, customer relationship management (CRM) and targeting (Karakostas, Kardaras, & Papathanassiou, 2005). Particularly, hotel industry is highly competitive considering that hotels generally offer homogeneous products and services, which contributes for hotels craving to distinguish themselves among competitors. To face this issue, guest satisfaction has become one of the crucial measures of hotel effectiveness and performance (Xiang *et al*., 2015). Guest satisfaction has been evaluated across time due to consumer surveys or focus group interviews. Although it is an efficient method, it often suffers from poor sampling and low response rates, producing vague assessments of guest experience and satisfaction. As such, the use of data analytics in hospitality management measuring guest satisfaction becomes an intriguing research question.

The unparalleled growth of Internet applications to travel and tourism has produced a surplus of new opportunities and challenges to consumers and practitioners. Travelers generally tend to conduct an online search about their preferred destinations in order to make their travel decisions. Afterwards they can read, and use the reviews as references to understand whether a place is their preferred destination.[1]

According to two surveys of more than 2000 American adults (Pang & Lee, 2008) among readers of online reviews of restaurants, hotels, and various services, between 73%

[1]These data was extracted from DoubleClick Inc. (2005). Search before the purchase: Understanding buyer search activity as it builds to online purchase. Accessed on 15[rd] of May 2015: http://www.innovation-marketing.at/news/newsmodul/upload/429181378_searchpurchase_0502.pdf

and 87% report that reviews had a significant influence on their purchase. Also, 32% have provided a rating on a product, service, or person via an online ratings system, and 30% have posted an online comment or review regarding a product or service. Furthermore, it is stated that consumers are willing to pay from 20% to 99% more for a 5-star-rated item than a 4 star-rated item.

Several studies have confirmed that the image of certain destination can be directly affected by travel-related websites (Choit *et al*., 2007; Govers & Go, 2005; Pudliner, 2007). In addition, Choi *et al*. (2007), claims that travel blogs will become a popular source for destination information. Furthermore, Pan *et al*. (2007), argues that the feedback that is available on travel blogs can be richer in content and more detailed than Likert-scale based survey questionnaires.

Stokes & Lomax (2002) argued that improved word-of-mouth communications can be an effective marketing strategy for small hospitality businesses. Also, Goldenberg *et al.* (2001), argue that customer's decision making process is strongly influenced by word-of-mouth. Since travel is typically experience related, opinions in online reviews should have a strong influence to travelers (Litvin *et al*., 2008). This affluence of consumer-generated data offers opportunities to make statistical inferences in order to understand consumer behavior in hospitality.

In hotel industry it is important to mention the growing market focus on what regards sustainability issues, namely responsible tourism, eco-hotels and sustainable friendly hotels, which led to a literature increase about these topics. The Telegraph newspaper per UK reporter Lola Pedro (2013)[2] suggests that in recent years sustainability in the travel industry has progressed from just a niche to an industry-wide priority. Some eco-conscious travelers no longer appreciate their luxurious hotels or lodgings being the focus of their holiday. Rather, these travelers are increasingly looking to the location itself to provide contribution to local communities and to be as unobtrusive as possible. It is stated that the travel industry contributes to a significant amount of pollution and waste. This fact conducts to the increasing traveler's interest in environmentally friendly holidays,

[2] Pedro, Lola. "Sustainable, responsible tourism: Luxury travel's new trend." Published on The telegraph on 15 Mar.2013 and accessed on 10rd December 2014: http://www.telegraph.co.uk/luxury/travel/1983/sustainable-responsible-tourism-luxury-travels-new-trend.html

as so does the number of innovative, luxurious and friendly experiences the hotels have to offer.

2.2 Business Intelligence applied to Hospitality

Nowadays, customers have at their disposal an unparalleled extent of power by which to share their brand experiences and opinions, regarding any product or service. The consumer voices can yield huge influence in shaping the opinions of other consumers. With this growing availability and popularity of opinion resources, new opportunities and challenges arise for hotels. Considering the loads of information available, the hotel industry is overloaded by data. Also, the majority of information is displayed in unstructured or semi structured way, being difficult to combine both qualitative with quantitative data within business analysis (Lau, Lee & Ho, 2005). Therefore, there is an urgent need for innovative techniques that can automatically extract business intelligence from large text collections and integrate fragmented information into BI databases.

Opinion mining and sentiment analysis techniques can perform those tasks, and automatically understand customer’s opinions and preferences. Allowing for the integration of text mining techniques as a means of information management, these can be used by hoteliers in a way to develop competitive and strategic intelligence, considering that TM is well-suited to several types of market intelligence applications (Pang & Lee, 2008).

There has been a growing tendency in using customer generated data to obtain insights into research problems that have not been well understood by conventional methods (Xiang *et al*., 2015). Data analytics allows the development of new knowledge and provides an understanding for support decision making in hospitality industry.

Hotels can develop innumerous benefits from adopting BI techniques, as extracting meaningful patterns and build predictive customer-relationship models from numeric data. Besides, it allows the environmental scanning of customer intelligence by analyze customer newsgroups, online boards, travel blogs, online customer surveys. Moreover, the improvement of efficiency of internal knowledge management by analyzing emails, patent databases and corporate documents, hoteliers may obtain up-to-date knowledge of potential travelers, thus gaining a competitive edge in running a lodging business (Lau, Lee & Ho, 2005). Accurate and timely information within customer intelligence enhances the hotel effectiveness and as a result customer satisfaction. Thus ignoring consumer sentiments might put companies in a competitive disadvantage and could also create significant brand image problems.

2.3 Sentiment analysis

The advent of Internet social media has triggered sentiment analysis, a recently developed web mining technique that can perform analysis on sentiments or opinions based on published reviews and comments. It aims to extract the text of written reviews for certain products or services by classifying them into positive or negative opinions according to the polarity of the review (Cambria *et al*., 2013). This classification method has been applied to a wide number of domains, including tracking political opinions, (Stieglitz & Dang-Xuan, 2013) product reviews (Nie *et al*., 2013), track sentiment trends (Glooret *et al*., 2009) and predict stock market trends (Wong, Xia, Xu, Wu, & Li, 2008).
Tourist destinations would naturally be one of the good application areas, as it is able to provide to tourists necessary information to purchase goods or services and to the hoteliers knowledge about the response from their customers and the performance of their competitors, having broader implications as a BI tool (Lau, Lee & Ho, 2005).

On the tourism and hospitality domains, sentiment classification has been studied recently by Ye *et al*. (2009) and by Shi & Li (2011). The former analyzed reviews from seven destinations using three machine learning algorithms (naïve Bayes, support vector machines and the character based N-gram model) achieving an accuracy of more than 80%. The latter study adopted support vector machine for analyzing an online published and prepared dataset, obtaining an accuracy of around 85% by using information based on TF-IDF (term frequency–inverse document frequency). Both works mentioned are focused on improving the accuracy of sentiment classification, not unveiling the products and services targeted by customers' reviews.

Another interesting study in this field, by Blair-Goldenshohn *et al.* (2008) used Google Maps data as input in order to analyze consumer sentiments towards hotels, department stores and restaurants. Polarity values (positive/negative and neutral) were used and the system was able to summarize sentiments regarding different aspects of the service provided, such as value for money, rooms, location, dining, and general service. This study provided a mechanism for further reducing the amount of data required to produce highly accurate sentiment classifications.

Sentiment analysis, or opinion mining, could have innumerous applications, highlighting the automatic upkeep of reviews and opinion aggregation websites, suggesting recommendation systems and improved human-computer interactions. Another possibility focus on bring up products/services ads when positive sentiments are detected, and more important, null the ads when relevant negative sentiments are discovered. This tool is an excellent method for extracting opinions from unstructured documents and handling many BI tasks, as perform trend prediction in sales or other relevant data. Also by tracking public opinions it contributes for reputation and brand management (e.g., Pang & Lee, 2005).

Despite being a tool with multiple advantages, sentiment analysis also have special challenges associated with it. One of them is over due to web text being classified as noisy (in lexical and syntactic levels). Word semantics in a particular review could contradict with the overall semantic direction of that review. The main solutions for this problem are the word sense disambiguation and inferring semantic orientation from association (Turney, 2002).

2.4 Text mining

The most advanced decision support and business intelligence systems usually incorporate machine learning techniques, mainly data mining for exploring patterns hidden in data that can be translated into useful knowledge (Witten & Frank, 2005). Text mining is a particular method of data mining that focus on analyzing the value of unstructured data such as raw text contained in a collection of instances of the problem being studied (Fan *et al*., 2006). These instances may be documents, comments, reviews or any other sort of related information, constituting the corpus for feeding text mining.

In other words, text mining explores data in text files to establish valuable patterns and rules and to point out trends and significant features about specific topics. Text mining tools convert unstructured data into structured data, reduce the dimensionality of data while keeping relevant information, and jointly analyze quantitative and qualitative data.

Some important authors regarding the use of this tool were Delen and Crossland (2008), which proposed the application of TM for analyzing literature and identify research trends. Their investigation aims to help researchers to conduct state of art reviews about a certain given subject. They argue that their TM approach can be valuable and applied to any research field. Their approach states that relevant words and terms are often extracted in order to build a corpus of knowledge among the input of data considered.

Also, in TM analysis sometimes searching individual words it is not enough, since many terms are composed of a sequence of words, called n-grams (e.g., “data mining”), when extracted from large texts it can be a valuable asset (Soper & Turel, 2012). Such method addresses the difficulties posed by single word semantics described in section 2.3.

Other interesting approach focuses on modeling a certain number of different topics, previously defined according to the distribution of different terms across the documents, this is achieved through LDA model (Blei, 2012). This model allows determining the probability of a chosen document belonging to each topic, grouping documents according to their proximity regarding the subject studied. This model can help identify which topics are capturing more attention and also to find gaps for future research.

Rosseti *et al*. (2015) followed this approach, using text mining for providing recommendation to specific users based on their previous online reviews. Based on the words frequency, they conducted topic modeling to make judged recommendations, also providing means to get feedback from users about the recommendations made. At the end of this article, the authors propose the usage of topic modeling for analyzing sentiment analysis as a possible future work to explore, which is precisely the problem addressed by the present thesis.

Another similar recent research published by Xiang *et al*. (2015) adopted a text mining approach for analyzing 60,648 reviews posted on Expedia.com. They used statistical methods for measuring word frequency to infer on guests’ experience and satisfaction. However, they consider single words only, not including multiple word terms, thus no n-gram analysis was conducted such as in the work of Soper & Turel (2012). Also, no clustering algorithms were incorporated, in order to gather the reviews in logical groups for providing managers with insights about the users’ feedbacks.

TM has innumerous applications, providing means for analyzing textual information contained in emails, customer surveys, online boards, corporate documents, patent databases. Broader industries such as tourism, financial services, insurance services, and manufacturing have successfully applied these techniques in their data-management functions and highlight the efficiency of this tool in managing huge amounts of text data. (Lau, Lee & Ho, 2005).Some of the factors contributing for the increased interest in this area according to Pang & Lee (2008), include the rise of machine learning methods, natural language processing and information retrieval, also the development of web-sites reviews aggregation and finally the increase in commercial and intelligence applications applied to business.

This process has improvements regarding the traditional survey approaches, as in data-mining data is self-revealed according to the owner's preferences and researchers set inquiries without regard to what kind of information is available. (Lau *et al*., 2005). It is important to note that text mining success not only depends on search hinge effectiveness but also requires the researcher to interpret the results precisely and with accuracy.

3. Material and methods

3.1 Data collection

For this empirical research, relevant information about a Portuguese eco-hotel from different data sources was collected, mainly the *Areias do Seixo*[3] hotel, located around 60 kms North of Lisbon, Portugal. Such information was materialized in comments written by customers and served as an input to the TM procedure.

A random set of comments was collected, during January and August 2015, the data comprised 401 different opinions, from the total 3179 reservations during this period, serving as the main input for the experimental procedure. The sample size is comparable to Pekar & Ou (2008) employing a sentiment analysis technique for evaluating 268 reviews of major hotels based on customer's reviews posted on the website "eopinions.com", using attributes such as food, room service, facilities, and price to automatically analyze customer sentiments towards those features.

The collection of comments was originated from six different sources, as shown on Table 1. Taking into consideration the dispersion of reviews about this hotel in both on-line and off-line domains, all the six sources were included.

Table 1 - Sources for the reviews analyzed

Nr.	Source	Nr. Comments	%
1	Trip Advisor	52.5	13%
2	Guest's book	272.5	68%
3	Follow up emails	40	10%
4	Evaluation website	4	1%
5	Direct emails	26	6%
6	Other	6	1%
	Total	**401**	**100%**

[3]http://www.areiasdoseixo.com/en/hotel-overview.html

One has to take into consideration that sometimes the same comment was written in two distinct data sources, namely Trip Advisor and Guest's book; for addressing this issue, the comment was attributed to half (0.5) for each of the two sources, justifying the decimals seen in Table 1.

Trip Advisor sampling

The applied corpus contains customer reviews of Trip Advisor, which is considered one of the most popular and well-known travel and vacation services microblog web-site (O'Connor, 2008). Into this platform customers can book, rank and review hotels, flights and restaurants. Customers can filter content based on rankings attributed by customers. These rankings are split into several categories like, value, rooms, location, cleanliness and sleep quality. The rating scale contains five values from "terrible" to "excellent", further extended to as 1 star to 5 stars. In this study, Trip Advisor rankings are not considered, once the main goal is to conclude on the sentiment classification regarding specific hotel terms and not an overall sentiment score for each review.

Guest's Book sampling

The applied corpus also contains customer reviews that follow a traditional approach, considering that Areias do Seixo adopts a different method regarding customer surveys, as it is not their policy to perform customer's questionnaires, thus they focus their strategy in other sources of information, namely the Guest's Book, created for guests to express their opinions about hotel products and services, upon to conclude on Hotel's quality.

This strategy makes customers write down their opinions when they leave the hotel and since they are not forced to (as the idea behind a normal questionnaire), it brings more spontaneity and will to write a review. Considering this as part of the Hotel's method to conclude on customer satisfaction, one took advantage of this source of information to incorporate it as a way to enrich the size of the information considered, even though it is an offline hand written source.

Follow up emails, Evaluation websites, Direct emails sampling

Although Trip Advisor is the most common and used source of information regarding customer opinions on travelling matters, other on-line sources of information were also included from different evaluation websites. This includes other important platforms beyond Trip Advisor, such as Zomato, which is a phone application used by customers to discover the best places to eat. Apart from websites, follow up emails (emails that hotel send to customer to ask for feedback) and emails sent directly and spontaneously from customers to Hotel, elucidating hotel managers on the strengths and weaknesses, were also included.

Hotel Director Initial Interview

In the beginning of this work, an unstructured interview took place with one of the hotel managers for obtaining his perceptions on the customer feedback, as well as to understand the main strategy followed by this particular eco-hotel, in order to provide in-depth knowledge for analyzing the results. This valuable in-loco knowledge allowed strengthening the discussion of the results.

3.2 Proposed approach

Figure 1 shows the approach proposed for extracting useful knowledge from the unstructured text contained in the customers' reviews. This approach is based on the work of Moro *et al*. (2015) for a literature analysis using a set of relevant articles on BI applications to the banking industry. Usually, TM involves two processes for building the corpus of reviews: cleaning of the text from irrelevant words such as articles and adverbs; and stemming, for reducing words to a single root word (e.g., "feelings" is reduced to "feel").

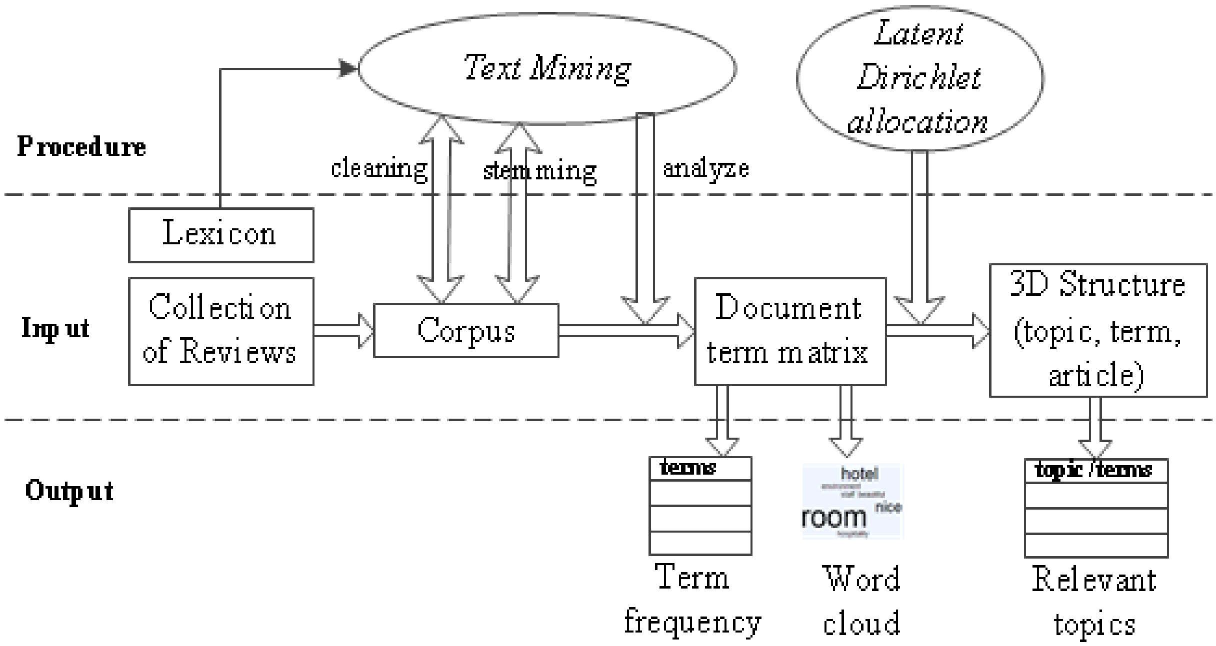

Figure 1 – Proposed approach

However, the present analysis focused specifically on sentiment analysis and hospitality issues; hence, besides the collection of reviews, which constituted the main input from where hidden patterns of knowledge were extracted, the procedure was also fed with a lexicon that established the dictionary of relevant terms for both sentiment analysis and hospitality. Therefore, the cleaning and stemming processes used the lexicon contained in the dictionary for reducing the reviews to sets of relevant terms.

The lexicon constituted from both dictionaries was compiled as a single input, for allowing a cross-domain relationship analysis between sentiments and hospitality. The main output from the text mining procedure is the document term matrix, as shown on Figure 1. This matrix has two dimensions: the reviews (usually text mining is performed over documents, hence the name) and each of the terms considered; each of the cells contains the frequency each term occurs in each of the reviews. For analysis, two user friendly outputs were provided: a table of frequencies, for counting the number of occurrences of each term, and a word cloud, for providing an easier visual interpretation of those occurrences.

Finally in order to generate the most relevant topics, the document term matrix served has an input for the latent Dirichlet allocation topic modeling. The latent Dirichlet allocation final output is a tridimensional matrix encompassing terms, reviews and topics. Therefore, for every topic it is possible to obtain a measure of its relationship to one of the dictionary terms through the β distribution. Also, for every review it is possible to check to which topic it suits better. The product of these three dimensions results in a very large structure. Considering the goal is analyzing the relation between sentiments and hospitality, only the most relevant sentiment and the most relevant hospitality issue were scrutinized.

3.3 Text mining for sentiment classification

This technique was used to evaluate 401 reviews of Areias do Seixo customers posted in different sources as a way to facilitate in producing organized information (which could also be used easily if we had bigger data, making scalable the approach undertaken). The usage of comments consisting in unstructured text uses the approach described has the advantage of allowing to include more than one source which, otherwise it would be difficult if the customers' input was provided in structured forms, implying that a common scale needed to be defined, which could render unfeasible such inputs. Also, free comments do not constrict customer's thoughts and ideas.

Besides TM, sentiment analysis approach was also applied. There are two methods in applying sentiment analysis technique: lexicon-based approach or corpus-based method. (Miao, Li & Zeng, 2010). Both require a pre-defined dictionary or corpus of subjective words. In this study, only lexicon-based approach was used, as it is the most common. However, in both methods it is necessary the implementation of a dictionary.

It makes sense to incorporate a dictionary that encompasses the both Hospitality and SA more common terms and concepts, rather than let TM algorithm to search, group and count words indiscriminately. Since the study focus on crossing two different areas "hospitality" and "sentiment analysis", for that reason it makes sense to construct a dictionary that act as knowledge base to associate keywords to specific concepts. Hence, two distinct dictionaries were implemented, one from SA domain and another from Hospitality domain, each of them containing a list of terms composed by one or more words (n-grams). Table 2 and 3 illustrate these two dictionaries.

In the process of constructing a dictionary a usual first step is to perform stemming, in order to reduce similar words to a unique term (e.g., “Hospitality”; “hotel”; “hotels”). In addition, comments were translated to a common language: English, to remove complications of Multilanguage comments. Moreover, abbreviate words were eliminated, to reduce the noisy environment, as well as stripping extra whitespaces and all words were converted into lowercase to facilitate term comparison by the TM procedure.

Automatically-coded lexicons have recently been developed, including sentiment-based lexicon (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011). However in this study, we used a manual approach extracting the dictionary keywords from different known hospitality and sentiment analysis authors.

Table 2 shows the dictionary for hospitality including all the reduced terms and several of the similar terms. The set of base terms was extracted from the dictionary published by Ingram (2003). Then, it was followed the approach from Lau *et al*. (2005), where important hotel attributes were also extracted for enriching the hospitality dictionary. Finally, an analysis of a sample of the 401 reviews allowed identifying additional terms, which were also included.

Table 2 - Dictionary for the hospitality domain

Reduced term	Similar terms or from the same domain *
decoration	decoration, décor, decorative, interior design, architecture, details, detail, interiors, designer, designed, decorated, decorating
different	different, difference, creativity, simplicity, boldness, unique, singular, innovator, innovation, creative, personalized, original, originality, genuine, special, alternative, authentic, individually
environment	environment, nature, sustainability, sustainable, ecological, environmental, green, environmentally, eco
eur	eur, euro, euros, money, expensive, cost, price, prices, pricing, credit card, quality/price
equipment	equipment, equipped, machines, machine, amenities, water, facilities, facility, property, building, fireplace, structure, materials, wood, roof, furniture, cement, chimney, tub, pool, shower, rocking chairs, air-conditioning
feelings	feelings, feel, feels, sense, senses, sensations, energies, energy
food	foods, food, restaurant, chef, chefs, taste, tastes, flavors, wine, wine experience, dishes, dish, cuisine, menu, meal, meals, entries, main course, dessert, breakfast, cook, cooked, cooking, cookers, vegetarian, gastronomy, order, ordered, bar, vegetables, fresh fruit, fresh fish, local products, herb, tea, cake, organic products, dinner, lunch, ingredients, drinks, vegetable garden, veg, vegan, minibar, horta, catering, buffet, recipe, pie, gourmet, pastries, garden, wine cellar, homemade, tasting menu, kitchen, sandwiches, salads, mojito, drink, eggs, bacon, dining service, cookies
friendly	friendly, friendliness ,kind, kindness, kindly, caring, nice, nicest, care, attentive, generosity, faithfully, friends, friend, empathy, sympathy, sympathetic, gently, welcoming, hospitable, pleasant, warmth
guests	guests, guest, clients, client, host, hosts, customer, costumers
holiday	holiday, holidays, vacation, break away, experience, activities, excursions, explore
hospitality	hospitality, accommodating, accommodation, hotel, hotels, resort, lodge, reception, front desk, maintenance, housekeeping, living room, rooms, room, check in, check-in, checked-in, checked out, bed, beds, bedroom, bedrooms, stay, villa, villas, residences, child travel bed, matress, terrace, private patio
location	location, locally, local, locale, place, spot, landscapes, view, views, sight, area, surrondings, sunsets, retreat, seclusion, scenery, sea, ocean, dunes, beach, lackhouse, beaches, waves, fire pits, fire circle, bonfire, swings, weather, surfing area, sea-side, seafront
people	people, humans, professional, employees, employee, staff, workers, team, co-workers, owners, personnel, waiters, house keepers, roomkeepers, administrators, trainee, receptionist, kitchen staff, waitresses
relax	to rest, relax, relaxing, relaxed, calm, quiet, massage, spa, jacuzzi, yoga, treatment, peace, zen, feel good, relaxation, chill
reserve	reserve, reserves, reservation, booking, availability, available, transport, book, booked, renting, shuttle
romance	romance, love, lovely, romance, romantic, passion, passionate, intimate, charming, sensivity, charmed, charm, magically, magical, magic, heavenly, poetry, honeymoon, ceremonious, wedding
site	site, sites, website, browsing, internet, connection, connections, wi-fi, computer, tripadvisor, magazine, opinion, review, email, mail, emailed, social networks, marketing, information, loads, facebook
tourism	tourism, travel, travelled, travelling, tour, touristic, trip, roundtrip, industry, visit, destination
trends	trends, trend, concept, theme, thematic, boutique, ethnic, ethnical, chic, exotic, hippie, style, lounge, trendy-friendly, zen-friendly, gay-friendly, anti-bullfighting friendly, friendly smoking

*** All terms are in lower case and separated by commas**

Some of the terms are constituted by more than one word, unlike the studies of Rosseti *et al*. (2015) and Xiang *et al*. (2015). By considering n-grams (Soper & Turel, 2012), the procedure can embed some context through the combination of a few words (e.g., "social networks").

Table 3 - Dictionary for the sentiment classification

Reduced term	Similar terms or from the same domain *
strong positive	brilliant, excellent ,fantastic, phenomenal, wonderful, superb, beautiful, spectacular, high-quality, top notch, top, incredible, best, greatest, delightful, memorable, very fresh, very nice, real change, perfect, delicious, amazing, happier, perfectly, very much, fascinating, gorgeous, unforgettable, speechless, dreamful, dream, great, fabulous, paradise, enchantment, exceptional, delights, outstanding, surprising, gorgeous, :-) ,deeply touched, extraordinary, 5 stars, 5 star, stunning, surrender, breathtaking, marvel, exclusive, delighted, splendid, pampering ,maximum, wow, happiness, precious, exceptional, dream place, thumbs up ,impressive, remarkably, adorable, overlooking, stunning, tastefully, enchanting, touching, amazingly, genuinely, favorite, exciting, magnanimous, impeccably, knowledgeable, loved, inspirational, perfection, beautifully, magnificent, stupendous, inspiring, blow-away, marvelous, dazzled, impressed, high standard, honored, super, love you, adores you
strong negative	terrible, awful, stupid, horrible, unfortunately, ridiculous, difficulty, really hard, too long, weaknesses, very bad, too slow, nightmare, stupidity, food poisoning, unnecessary, extremely slow, nuisance, regrettable, not respect, disturb, discrimination, prejudice, worse, price disparity, discrimination policy, angry, cheated, not worth, not professional, lie, overpriced, unnoticed, envy, endless, criticism, not recommend, complaint, disappointed, unpleasant, nauseating, disaster, uncomfortable
ordinary positive	cool, good, fashionable, better, cozy, modern, goodness, helpful, peaceful, beauty, quality, warm, respect, tasty, recommend, spacious, pleasure, elegant, sincere, joy, enjoy, enjoyed, smiles, serenity, worthy, comfortable, tasteful, excellence, come back, wealthy, harmony, sanctuary, happy, success, encouraging, luxury, sublime, affection, elegance, merit, diversity, fine, gratitude, inspiration, smiles, finely, picturesque, congratulations, congrats, thanks, thank you, remember, well-tasted, affection, surprise, healthy, large, fancy, interesting, grateful, courage, thankful, celebrate, cute, curious, delicacy, appreciate, return, like, liked, enthusiasm, sophisticated, enthusiastic, exclusivity
ordinary negative	bad, nervous, loss, aversion, sad, difficulty, quite small, little scattered, expensive, shame, unbalanced, spoiled, burst, tears, shy, back pain, missed, raw, blood, ignorance, bad time, mistakes, tough, failed, stressful, against, regret, doubly, stop, noise, rivals, hard, lacks, limited choice, less, dirty, blamed, concern, problems, trouble, tiny, missing, apology

*** All terms are in lower case and separated by commas**

In table 3, the sentiment classification dictionary is exposed. For defining this dictionary, first an accurate classifier (scale) was required for compiling indicators of sentiment; then the sentiment is determined by comparing comments against the expert-defined entry in the dictionary, which makes it easier to determine the polarity of a specific set of words. The scale used in order to construct the sentiment analysis dictionary following the approach of Hu *et al*. (2012). Thereafter, the dictionary was enriched with terms representing sentiment intensifiers, following different polarity, including "strong positive";" ordinary positive"; "ordinary negative"; and "strong negative" categories.

In addition, as the definition of a dictionary and grouping terms under a unique reduced term are subjective, in order to reduce such subjectivity, the author had the help of a Marketing and Hospitality specialist to analyze all decisions, and also an information system specialist for validating that the dictionary was compliant with the TM tool used, to further extend the validation of the dictionaries and TM procedures.

3.4 Text mining tools

To perform the TM review, we chose to use the R statistical tool, which is open source and provides flexibility through the installation of packages published by a large number of supporters in the CRAN (Comprehensive R Archive Network - https://cran.r-project.org/). For the text mining procedures, the "tm" package was adopted, considering it was specifically developed for conducting the text mining functions needed to analyze text (Meyer *et al*., 2008). This package provides functions for converting unstructured into structured data, reducing dimensionality of data while keeping relevant information, and jointly analyzing quantitative and qualitative data.

To gather the topics which group comments, the "topicmodels" was adopted, for it receives as input the data structures produced by the "tm" package in order to provide basic infrastructure for fitting topic models, (Grun & Hornik, 2011),within the "topicmodels" package, the LDA algorithm is implemented (Blei *et al*., 2003).

Latent Dirichlet allocation (LDA), is a three-level hierarchical Bayesian modeling process that groups collections of items in topics defined by identified words or terms and the probabilities that each of them characterizes the topic (Blei, 2012).Such model enables analyzing the relative relevance of each term using the β distribution value, which characterizes the relation between the topic and the given term. All β value are negative, thus to facilitate the interpretation, the absolute value for all cases are considered. A β closer to zero represents a stronger relation between a term and its corresponding topic.

Within the topic models package, the LDA algorithm is implemented and it can be computed with only two parameters, the desired number of topics and the document term matrix created for the TM.

3.5 Classification of topics

As observed in the previous sections, the intersection of different domains allows finding existing relations for the categories considered. In this section topics are classified through LDA topic model in order to find those relations. Since both dictionaries are merged, this could mean that a topic could be better characterized by one of the terms relating to a single category. Nevertheless, this technique provides interesting insights for relations between categories (Moro *et al*., 2015).

In this paper, only the most probable topic according to LDA for a given comment was considered. To proceed with LDA computation, the number of topics is a required parameter which can be tuned for optimal results (Yi & Allan, 2009). Following the approach of Delen & Crossland (2008) the number of topics was set to half of the terms considered in a first stage, so initially 12 topics were considered nevertheless experiments for tuning this number determined that 9 topics provided the most solid-ground results.

Such decision was based on analyzing the output consisting in the topics and their characterizing terms: with a number of topics equal to 12, some of the topics overlapped in the most relevant terms (although with different β values) with those being distinct in the remaining terms. Therefore, we ran the LDA procedure in a few iterations while reducing the number of topics for tuning the results (Blei, 2012).

4. Results and discussion

The results were achieved through the process described, with the help of the dictionaries, where unorganized information text was translated into meaningful data structures, which were used for data mining purposes. Thus, the qualitative text data was combined with the quantitative data, through computer programs, for a more comprehensive business analysis into the hotel industry and converging on an important tool for managerial decision making.

The results are presented in two different sections: in the first, results are analyzed based on the term frequencies for the whole 401 customer's reviews collected, through text mining technique. The respective results are shown using a table and a word cloud, which uses a larger font size for the most frequent terms. After this global analysis, the results are presented in topics generated by LDA, which groups similar comments into topics. In the second section, examples of comments are displayed by topic in order to illustrate this trend.

4.1 Exploratory Data analysis

The results obtained for the text mining procedure, presented in Table 4, exhibits the number of occurrences for each of the reduced terms according to the equivalences determined from the dictionaries (Table 2 and 3).

Table 4 –Term frequency for the sentiment analysis applied to Hospitality domain

#	Term	Frequency
1.	strong positive	739
2.	ordinary positive	601
3.	food	445
4.	hospitality	424
5.	location	290
6.	romance	230
7.	people	150
8.	different	133
9.	decoration	110
10.	relax	107
11.	holiday	89
12.	ordinary negative	68
13.	equipment	56
14.	environment	53
15.	feelings	51
16.	strong negative	46
17.	site	43
18.	eur	32
19.	trends	29
20.	friendly	28
21.	tourism	28
22.	reserve	26
23.	guests	9

Note: sentiment classification terms are grayed for easier identification

The grayed rows represent terms related to the four types of sentiments used for classifying customers' satisfaction. It is notorious the fact that positive sentiments are ten times more frequent than their respective negative counterparts. Moreover, the text mining procedure recognized a perceived strong positive sentiment as the most frequent term of both sentiment and hospitality domains, occurring 739 times in the whole 401 reviews, with an ordinary positive sentiment being the second most frequent term, with 601 occurrences. This result is aligned with the hotel manager's perceptions, who stated in the previous discussion that the majority of customer's verbally expressed that they were definitely happy and delightful with their stay, with the reviews confirming this fact.

Figure 2 - Word cloud for Hospitality domain

Figure 2 shows the word cloud for terms from the hospitality domain only, providing a visual interpretation of the results. The sentiments were excluded since only four sentiment classifications are considered to allow a clearer picture of the relevance of hospitality terms.

First, it should be stressed that the term "hospitality" is accounting for accommodation related terms only, as shown on Table 2. The global results, presented on both Table 4 and Figure 2, with a total of nineteen hospitality terms, show that the words "food" and "hospitality" are clearly the main hotel attributes mentioned by the customers; particularly in measuring the main reasons for customer satisfaction in this Eco-Hotel as a tourism destination. There is also a relevant interest in location, romance and people.

The second level analysis is focused on the use of LDA to generate and group the topics which were parameterized into 9, as shown in Table 5. This table will be more interesting for this study, since it allows relating the sentiment expressed by customers to hospitality attributes considered, thus identifying how strong this relation is and what are the most common customer's sentiments regarding this Portuguese Eco-Hotel.

Each topic is presented in horizontal lines, with the column labeled "Hospitality term" presenting the most relevant hotel attributes and the column labeled "sentiment term" the most relevant sentiment regarding each topic, and also column "β" disclosure β distribution values in respect to a given topic (where a smaller value represents a stronger relation). The number of comment column (#) presents the number of comments/reviews within each topic. For each topic, there is always a dominant term, with a β value that matches it closer to a certain sentiment term. Given that the two most relevant terms are shown for each topic, one regards a hospitality term, the other regarding a sentiment term. This enables to analyze each topic as a sentiment expressed for a hospitality attribute.

Table 5 – Topics discovered for the sentiment analysis applied to Hospitality

Topics	#	Hospitality Term	β	Sentiment Term	β
1.	65	food	0.53	ordinary positive	3.01
2.	58	location	2.24	strong positive	1.24
3.	52	different	1.84	ordinary positive	1.32
4.	45	romance	2.07	strong positive	1.11
5.	39	food	1.92	ordinary positive	1.14
6.	39	location	1.5	strong positive	1.57
7.	37	hospitality	1.91	ordinary positive	1.44
8.	35	site	2.42	strong positive	1.66
9.	31	hospitality	2.06	strong positive	1.18

is the number of comments per topic; β correspond to the correlation between the term and the topic

The most noticeable characteristic of all the topics unveiled is that all of them are related to positive feelings, with four of them representing an “ordinary positive” sentiment, and the five remaining representing a “strong positive” sentiment. This is a confirmation of the text mining results achieved on Section 4.1 a.

First, it is remarkable the fact that in general there is not such a high level of difference between the β values for hospitality and sentiment terms within each topic, by comparison with the results of Moro *et al*. (2015). The largest difference happens for the first topic (0.53-3.01) while the remaining topics show consistent β values where the lower β values is above half the β value for the other domain lower term. By comparison, Moro *et al*. (2015) show results with consistently large differences for the most relevant term to the second most relevant (the largest being 0.03-4.35). Such result reduces the problem identified in their study about weak correlation terms, strengthening the relations discovered between the two distinct domains.

The first topic, being the most mentioned service regarding hospitality terms, is best identify with “food” gets a matching of 65 comments, having a significant lower β value (0.53) meaning that the relation between the topic and the hospitality attribute is strong. However the sentiment term associated presents a higher (3.01), resulting in a distant relation from the topic. Despite the given β value “ordinary positive” is still the most common sentiment concerning those 65 customer reviews, on what regards food.

In order to confirm this hypothesis, topic number five, also underlines the sentiment “ordinary positive” and its relation with “food”, getting a closer β value in this topic, 1.92 on what regards “food” and 1.14 for the sentiment associated, reinforcing this relationship, enclosing a total of 104 reviews from the universe of 401 analyzed, including the first topic.

Both the first and fifth topics represent interesting discoveries, and where one can hypothesize that by targeting customers with attractive food in the service offered may also serve the purpose of retaining them by transforming ordinary to strong positive feelings. These are expected results, considering first that food is highly related with satisfaction (Namkung & Jang, 2007), and second that the food industry employs a large number of workers, stating the difficult challenging of managing them in order to fully satisfy customers (Koys, 2003).

The second topic, which best identifies with “location” gets 58 matching comments, having a lower β value (2.24), when comparing to the sentiment term associated, with a 1.24 β value, considering the “location” of this hotel as a “strong positive” sentiment expressed by Areias do Seixo customers. This puts emphasis on the number of positive adjectives expressed by customers associated with its location, within just five minutes’ walk to the beach, far enough from largely populated cities, but reachable in a 40 minutes’ drive from Lisbon, where the main airport is located, as well as many of the Portuguese national tourists live.[4]

Topic number six also emphasizes this “strong positive” sentiment regarding “location”, presenting very close β values, meaning a tighter relation between both terms and the topic. These are expected results concerning Kandampully & Suhartanto (2000) study, focusing on hotel’s location as being one dimension of hotel image attributes, and as a consequence one of the most important factors considering customer intention to repurchase, recommend and exhibit loyalty.

Another important relation is mentioned in third topic, associated with “different”, expressing an “ordinary positive” sentiment, and one can notice the strength of this relation, highlighted on Table 5. The Hotel is considered to be a “thematic luxury Eco-Hotel”, focusing its strategy on differentiation, and customers seem to be pleased with that hospitality attribute. Nevertheless, the closer relation of both terms implies that hotel manager’s may still have room to improve and transform this ordinary in strong positive feelings.

The forth topic is “romance”, gathering 45 comments, and also presenting a strong relation between the both terms and the topic. It shows a β value of 2.07 for romance, and a closer value to the sentiment term expressed (1.11); such a result translates that customers express a “strong positive” sentiment regarding romantic characteristics and details of the place. This represents a good investment from the hotel manager’s considering the hotel’s web site presents several offers and vouchers related to romance.[5]

[4]These data were extracted from Areias do Seixo website on 3rd of December, 2014: http://www.areiasdoseixo.com/hotel-directions.html

[5]These data was extracted from Areias do Seixo web site on 15rd of December, 2014: http://www.areiasdoseixo.com/hotel-products.html

Topic number 7 stresses "hospitality" and its relation with an "ordinary positive" sentiment. This is a strong relation, looking at its approximate β values, 1.91 for "hospitality" and 1.44 for the sentiment combined. However, topic number 9, despite taking into consideration the same "hospitality" term, presents a different sentiment (with 37 different customer opinions) where the sentiment emphasized in this case corresponds to "strong positive", highlighting its lower β values, meaning that this is a strong relationship. This relation with the topic "hospitality" means that customer value the quality of the hotel amenities and its service and its one of customer satisfaction factors in this hotel, but it still leaves room from improvements.

It should also be stressed the interesting result displayed by topic number 8, with the hospitality term "site" getting a match of 35 comments, where customers express a "strong positive" sentiment. According to the Hotel manager in an initial interview, one of the main reasons for people to choose this hotel was to see a picture of the place in the Internet or in an International magazine, read an opinion of an influence travel blogger or on Trip Advisor and also searched for Areias do Seixo website. Despite being one of the main reasons for people to choose this place, it is also one of the main hospitality products contributing for a higher customer satisfaction in Areias do Seixo hotel, accordingly to Results shown in Table 5.

One consideration can be made, from the nine topics, location, hospitality and food are the terms which best identifies with six of them, totalizing 269 comments from the total of 401 (around 67%), making of location, hospitality and food the main valued hospitality terms mentioned by Areias do Seixo customers. Figure 3 provides a visual interpretation of the results exposed on Table 5, with a larger line for the topics with more comments.

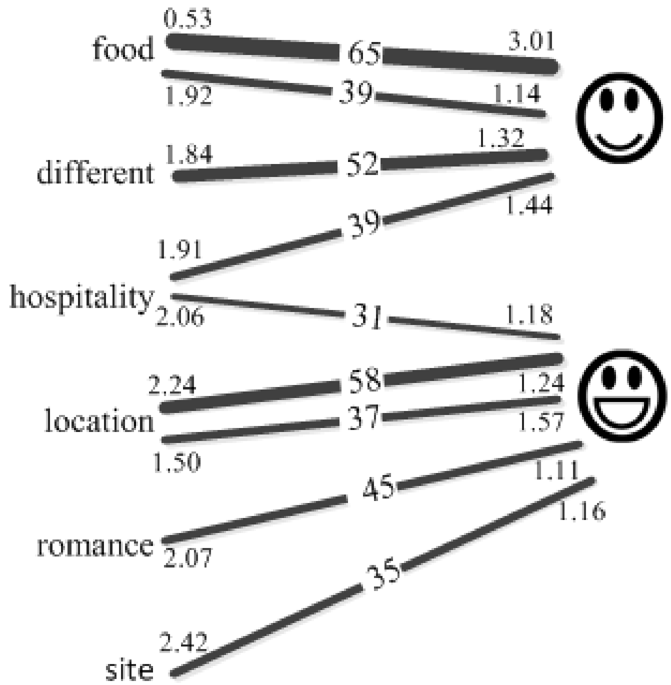

Figure 3 –Topics discovered in a picture

Although results don't show an emphasis on "people" neither on "decoration", this is an unexpected result since it is considered by Hotel Director as one of the main reasons for people to visit the place. Furthermore, Areias do Seixo is considered to be a Green Hotel, making great efforts in environmental issues and initiatives, however in our results this environmental awareness is not mentioned, this is an unexpected result, as the Hotel is considered a known and respected Eco-Hotel. According to Richard Hammond, founder of Green Traveller, "sustainability is still not a top criteria for choosing a holiday destination, things such as location, price and facilities are still the main drivers, however, being green has established a secondary consideration adding value to final customer, and making this a growing market tendency".

Finally, in general the processed results state that customers appreciate the place, for its unique location, the quality of its amenities, romantic surroundings and characteristics of the place, as for the hotel's positive reviews and Internet visibility. However, it is notable that services and attributes as food services, hotel amenities and difference as an hotel attribute still have room for improvements, as management recommendations. These results are valuable for hospitality management, supporting decision making for improving the value perceived by customers, validating the approach proposed.

4.2 Validation of LDA topics

In the previous section, LDA was applied in order to group comments under disclosed topics, characterized by the respective terms, as shown in Table 5, suggesting the sentiment expressed under the major hotel attributes considered. However, this approach, as it is automatic, contains certain limitations according to Thomas, McNaught, & Ananiadou (2011), such as document clustering to be completely dependent on the technique used for creating the clusters, which is based on term identification, so the problem here consists in some terms having different meanings based on the remaining text, where subjectivity detection can be a difficult task.

In this section, this issue is addressed by identifying and analyzing one representative comment for each topic (see table 6). This approach was based on full text manual analysis of the chosen 9 comments, in order to confirm the hypotheses suggested by the topics found, in a similar approach of the proposed by Moro *et al*. (2015). The numbering of topics is the same as for Table 5, while the column “frequency” stands for the number of occurrences of both the hospitality and sentiment terms, considering also the dictionaries expressed in Tables 2 and 3.

Table 6 -Sample of representative comment per topic

Topic	Review	Sentiment	Hospitality term	Frequency	Source
1	Breakfast: The choices were wonderful and it was like being in someone's kitchen helping ourselves to food. Dinner: We had tasting menu. All dishes were so creative and tasty. We also like the way each chef serving their own dishes to tables rather than waiters going back and forth the kitchen and floor. It added to the taste that we could see the chefs cooking in their open kitchen. We are writing to give you our feedback on our stay at your hotel. We stayed in one of your villas facing the sea for 4 nights. It was our first time to Portugal and we can't tell you enough how amazing the whole trip was. We always travel all over Asia, Europe and USA every year based in London and we can confidently say this was one of the best experiences we've had. Staff (receptionist, kitchen staff, waiters, waitresses, house keepers): All so extremely personal, friendly, smiley and helpful. Villa: The design was at its best and the cleanness of the villa was second to none. Towels were all soft as new and the bed linen was so white and crisp with homely duvet. The design of the villa was practical and beautiful. The decorating style was so inspirational. We loved using the well equipped kitchen with bbq facility outside. The kitchen was all immaculate, didn't have any trace of other people who stayed and used the facility. Garden: It was natural and loved the veg garden. We bumped into one of the staff who was picking vegetables for the kitchen and she kindly offered us some herbs for our bbq and even sent someone to our villa to give us a box of garden vegetables. We were so grateful and happy. When we went back to our villa after dinner, our beds were turned beautifully, candles were lit everywhere. We can not wait to go back to the hotel again and we do hope the quality of things we mention above will be kept well. We thank you so much for looking after us so well and we are grateful for the sweetest memories you gave us. Look forward to seeing you again.	☺	Food	32	3
2	This was the perfect place to spend our honeymoon. Feels like paradise. Peaceful, beautiful, the staff is amazing. A 5 star experience! Thank you!!! Thank you	😃	Location	6	2
3	Food very original and genuine, healthy and tasteful. I recommend the spa for the quality of its treatments. Unique hotel, with a unexpected location, across the dunes and the sea. An extreme well-tasted place, relax service and well personalized, bold design, fitting perfectly in the nature. Very spacious rooms, with a cozy décor and dreamful bathrooms. A magnify place to be in a romantic environment, with details full of charm. It was a magical stay, very special. To come back one day.	☺	Different	10	2;3
4	Dear two Joanna's :-), Sergio, Carina, Maria and Philip . How are you? We had a safe trip back home but we miss you and your lovely place very much. As we promised you find in the attachment two pictures of the paintings we bought of our "home" woman-artist. We are so happy with them and they fit into our apartments perfectly!:-) We want to say a hearty thank you again for your kindness, your cordiality and for your excellent service! We will definitely come back soon and perhaps even for our marriage;-) We saw the beautiful scenery you made for the American couple and that really left us speechless. Lovely greetings from us and please keep in touch!	😃	Romance	8	5

5	Thank you for such a great week for me and my group at Yoga Retreat. We love everything, the dunes and the bonfire whit Philip and Lucas. The dinner at the fish restaurant with Martha. Love that she stay with us. We love the warm girl in the spa Martha, and all kindness from Julia and all staff.	☺	Food	4	2
6	We stayed here for a short stay after a trip to Lisbon. In the future we'll certainly stay longer. The location and grounds were perfect, the rooms even better and finally the staff still beating both of those. The restaurant opened us to new foods and ways to present them. The drinks were fantastic and the beach was beautiful. The walk from the hotel to the ocean was through dunes with colorful vegetation ending with a tremendous view of the waves. We hope to make it a regular place to stay in the future.	😃	Location	10	1
7	Dear Daniella. Just back home after a very nice roundtrip in north and central Portugal. We specially want to thank you for our fantastic stay in your hotel. Next time we really hope to stay longer in Santa Cruz beach. We appreciate the perfect reception upon our arrival and also your kindly and lovely smile! After travelling the world round finally we discovered that heaven exist (but only) on earth and in Portugal haha. The hotel was really " unhavre de paix " and a paradise. Be sure that we will inform our colleagues specialized in " Vip Travel " in Belgium. Once again 1000 x thanks.	☺	Hospitality	8	5
8	We had a fantastic stay at Areias do Seixo. You have a very wonderful hotel with a great environment, delicious food and amazing rooms. The whole nature environment is also a very calm and beautiful experience. The only thing that definitely needs to be improved is the extremely slow Wi-Fi. Some website never loads and that became a very stressful situation and I think totally unnecessary. Today you expect it to just work fine. The small Navio restaurant even had a better connection and I understood when I was able to use one of the computers in the reception that you also had another faster connection.	😃	site	11	3
9	We have seen and experienced many hotels but what the Areias do Seixo offers surpasses everything. An incredibly exciting, exciting and wonderful place. An architecture and interiors which the guest can not rest and this I mean the positive sense. There is always something new to discover. One can place in different corners and amazed. The team is open, friendly, warm, knowledgeable, funny, erected, restrained and makes the hotel what it is. A DREAM! The restaurant is maybe for some guests specifically but it has its own style and fits finally to the hotel. The quality is excellent! The rooms are amazing and yes, enter the 1st time remains open mouth. The size and the architecture of the room and the bathroom are impressive. We will definitely recommend the hotel and would definitely come again. THANK YOU FOR THE DREAM Ok, it is rather expensive and the selection rather modest but the hotel itself makes up all this.	😃	hospitality	15	1

Source: 1 - Trip Advisor; 2- Guest's book; 3- Follow up email; 4 - Evaluation website; 5- Direct emails; 6- Other

In order to select the most relevant comments per topic two metrics were considered: the number of different terms mentioned in each comment, and the total number of times each of the sentiment term and hospitality term occurred. As an example, topic 3 groups a total of 52 comments, for each of those comments, the number of occurrences for the respective “sentiment term” and “hospitality term” were extracted from the document term matrix, where the one with the higher frequency of terms was selected. In this case the selected comment presents a frequency of 12, with 3 “different” terms, and 9 “ordinary positive” terms. This reinforces the relation between the different aspects of the entire hotel as an “ordinary positive” sentiment accordingly to customers.

Topic one best matches with the example above, illustrated in table 6, where “food” is the most relevant term, showed by the number of food terms mentioned (23), in contrast with only (9) regarding sentiment term. This explains the difference in β values, as mentioned before. This comment confirms the importance that food has on customer satisfaction, and therefore on sentiment classification has a positive attribute in hospitality.

By looking for topic 8, one can notice that it clearly shows a weakness of this approach, although it groups 35 different comments, the relation between the sentiment and the hospitality terms is quite distant. The comment used as example (being the comment having the larger number of occurrences regarding both terms) shows an unpleasant situation regarding “site” hospitality term. When analyzing the entire comment it couriers a “strong positive” sentiment due to overall hotel characteristics, nevertheless when converging specifically on “site” terms it is difficult to associate the presence of this strong sentiment regarding the corresponding topic. The client considers the existence of an extremely slow wi-fi as a negative sentiment, as being a “stressful situation” and “totally unnecessary”. This topic reflects one of the major limitations of similar automated approach based on clustering methods.

According to Kumar & Sahoo, 2013, “clustering aims at finding a subset of items which are more similar than others using similarity measures”. However, clustering algorithms, as the case of LDA, may contain a problem, implying to be inevitable the existence of comments that cannot match any of the given topics, leading to issues such as the one with topic 8, illustrated in figure 4.

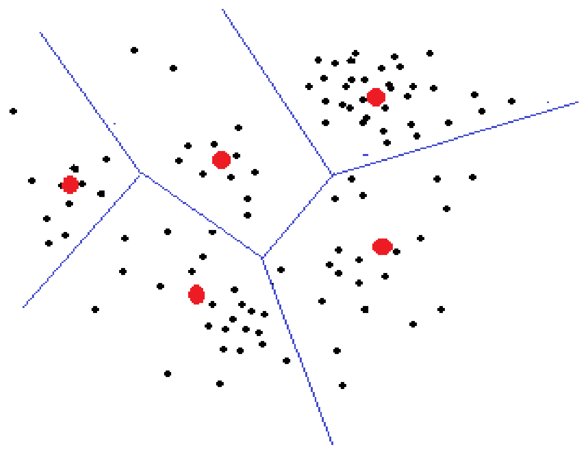

Figure 4 - Clustering algorithm

5. Conclusions

While text mining has been used as a new paradigm in many areas, very few applications, that fully explore its capabilities, were found in the hospitality field.
In this paper, sentiments polarity through text mining techniques of more than 400 customer's reviews were applied, as well as the identification of inherent relationships, using LDA modelling, between two domains of variables in hotel industry: sentiment classification and hospitality issues from the chosen Eco-Hotel.

The uniqueness of this study relies on the use of unstructured data from several sources to understand customer perceptions and feelings of a single hotel on a scale that was not available through traditional guest surveys studies. Also, the present research is justified by the work of Rosseti *et al*. (2015) considering their recommendations for future work focus on the usage of topic modelling for analyzing sentiment classification. As so, the present work fills such void in hospitality research.

This study contributes to the literature in several ways. Firstly, it provides a scalable sentiment analysis process applied to a specific hotel unit, which is a fundamental contribute to Marketing strategy, namely CRM, becoming a fundamental process in order for hoteliers to increase competitive advantage and create intelligent customer databases.
Also, the usage of latent Dirichlet allocation topic modelling for discovering costumer feelings generated by several hotel issues when crossing both semantics of sentiment polarity and hotel domain is a major contribution of this study, considering that the novel trends and generalized opinions unveiled may be used in order to improve hospitality business. Another contribution lies on, the proposed method be applied to an eco-hotel unit where topics find may expose how guest' satisfaction is being perceived, hence providing valuable knowledge for hotel managers' to understand the strengths and weaknesses of a specific unit.

From a practical point of view, this study enforces that core sentiments expressed through the mainstream hotel issues are deeply strong positive and ordinary positive. Customer retention seems to be associated with targeting, justifying customer satisfaction by the correlation between the unexpected location, the quality of its amenities, romantic characteristics of the place as for the hotel's positive reviews and Internet visibility.

Although, as management recommendations, food services may need to be improved, as well as the hotel' amenities and its differentiation focus strategy. These results are valuable for hospitality management, validating the approach proposed.

Nevertheless, the present study comprises several limitations and the findings should be interpreted with caution. First, no negative sentiment characterized topic was found, limiting the value to enhance the satisfaction perceived by customers, even though negative reviews are a less representative subset, according to topics found. Furthermore, considering that the sample represents only a single hotel unit in Portugal, the specific hotel issues identified in customer reviews obviously reflects the perceptions of location-related aspects of this hotel. Sentiment classification and guest satisfaction could be considerably different in another cultural context. Another limitation relies on clustering algorithms, as some reviews cannot match any topics, this issue was addressed by minutely analyzed the most relevant comment for each topic with the purpose of confirm the given results. Nonetheless, these potential limitations does not reduce the internal validity of data and thus not harm the purpose of demonstrating the power of sentiment mining techniques in the field of hospitality.

Future research may consider applying a fully automated system approach, as this proposal is a hybrid method containing the efforts of computer programs and manual labor. The ideal option should aggregate both in a single system as a technological development.

Overall, taking into consideration the actual globalization phenomenon and the development of new technological systems applied to management, this research makes us be facing, not just an academic case study, but a practical example to follow in the future, for the development of an innovative methodology that can conduct many companies through a remarkable marketing strategy, characterized by customer focus and competitive intelligence. As such, it is expected that this study sets an example for the development of business intelligence systems applied to hospitality marketing and management.

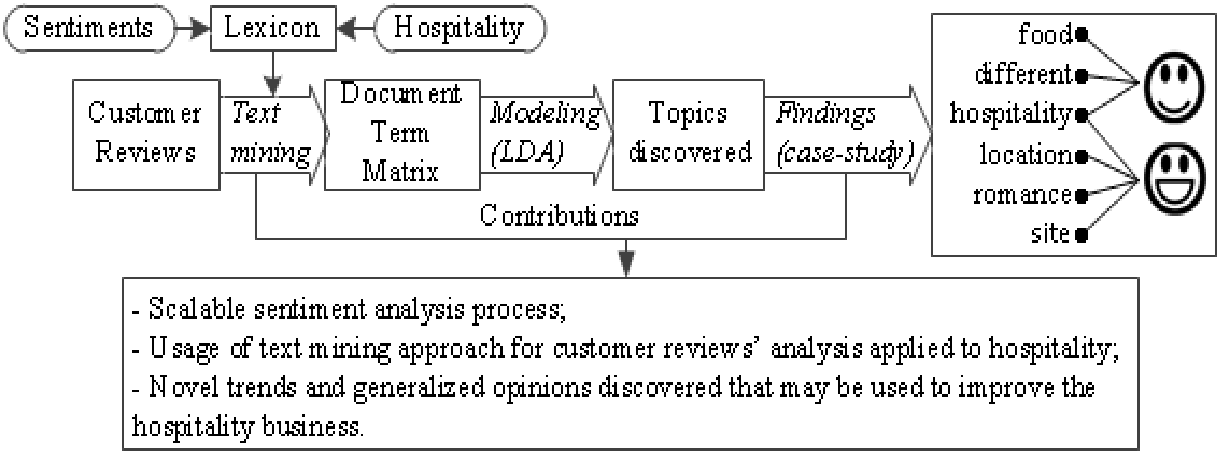

Figure 5 - Graphical summary

Bibliography

Articles

Bai, X. 2011.Predicting consumer sentiments from online text. ***Decision Support Systems,*** 50(4): 732–742

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G. A., & Reynar, J. 2008. Building a sentiment summarizer for local service reviews. ***WWW Workshop on NLP in the Information Explosion Era***, vol.14.

Blei, D. M., Ng, A. Y., & Jordan, M. I. 2003. Latent dirichlet allocation. ***the Journal of machine Learning research***, 3: 993–1022.

Blei, D. M. 2012. Probabilistic topic models. ***Communications of the ACM***, 55(4): 77-84.doi:10.1145/2133806.2133826

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. 2013. New avenues in opinion mining and sentiment analysis. ***IEEE Intelligent Systems***, (2): 15-21.doi:10.1109/MIS.2013.30

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.

Delen, D., & Crossland, M. D. 2008. Seeding the survey and analysis of research literature with text mining. ***Expert Systems With Applications,*** 34(3): 1707–1720.

Fan, W., Wallace, L., Rich, S., & Zhang, Z. 2006.Tapping the power of text mining. ***Communications of the ACM***, 49(9): 76-82. doi:10.1145/1151030.1151032.

Gloor, P., Krauss, J., Nann, S., Fischbach, K., & Schoder, D. 2009. Web science 2.0: Identifying trends through semantic social network analysis. ***Computational Science and Engineering, 2009.CSE'09. International Conference***, vol. 4: 215-222.IEEE

Goldenberg, J., Libai, B., & Muller, E. 2001. Talk of the network: A complex systems look at the underlying process of word-of-mouth. ***Marketing letters***, *12*(3): 211-223.

Govers, R., Go, F. M., & Kumar, K. 2005. Virtual Tourism Destination Image Innovating measurement methodologies. ***III International Doctoral Tourism and Leisure Colloquium***.

Hornik, K., & Grün, B. 2011.Topicmodels: An r package for fitting topic models. ***Journal of Statistical Software***, 40(13): 1–30.

Hu, N., Bose, I, Koh, N., & Liu, L. 2012. Manipulation of online reviews: An analysis of ratings, readability, and sentiments. ***Decision Support Systems***, 52(3): 674–684.doi:10.1016/j.dss.2011.11.002

Ingram, H. 2003. Dictionary of Travel, Tourism & Hospitality (3rd ed.), ***International Journal of Contemporary Hospitality Management***, 15(7):413-414. doi:10.1108/09596110310496079

Karakostas, B., Kardaras, D., & Papathanassiou, E. 2005. The state of CRM adoption by the financial services in the UK: an empirical investigation. ***Information & Management,*** 42(6): 853-863. doi:10.1016/j.im.2004.08.006

Koys, D. J. 2003. How the achievement of human-resources goals drives restaurant performance. ***The Cornell Hotel and Restaurant Administration Quarterly***, *44*(1): 17-24.

Lau, K, N., Lee, K, H., Ho, Y. 2005. Text Mining for the Hotel Industry. ***Cornell Hotel and Restaurant Administration Quarterly***, 46(3): 344-362.doi:10.1177/0010880405275966

Levy, Y., & Ellis, T. J. 2006.A systems approach to conduct an effective literature review in support of information systems research. ***Informing Science: International Journal of an Emerging Transdiscipline***, 9(1): 181-212.

Litvin, S. W., Goldsmith, R. E., & Pan, B. 2008. Electronic word-of-mouth in hospitality and tourism management. ***Tourism management***, 29(3): 458-468.

Meyer, D., Hornik, K., & Feinerer, I. 2008. Text mining infrastructure in R. ***Journal of Statistical Software***, *25*(5): 1-54.

Miao, Q., Li, Q., & Zeng, D. 2010. Fine-grained opinion mining by integrating multiple review sources. ***Journal of the American Society for Information Science and Technology,*** 61(11): 2288–2299.

Moro, S., Cortez, P., & Rita, P. 2015. Business intelligence in banking: a literature analysis from 2002 to 2013 using text mining and latent dirichlet allocation. ***Expert Systems with Applications***, 42: 1314–1324.doi:10.1016/j.eswa.2014.09.024

Mostafa, M. M. 2013. More than words: Social networks' text mining for consumer brand sentiments. ***Expert Systems with Applications***, 40, (10): 4241-4251.

Namkung, Y., & Jang, S. 2007. Does food quality really matter in restaurants? Its impact on customer satisfaction and behavioral intentions. ***Journal of Hospitality & Tourism Research,*** *31*(3): 387-409.

Nie, X., Liu, L., Wang, H., Song, W., & Lu, J. 2013. The Opinion Mining Based on Fuzzy Domain Sentiment Ontology Tree for Product Reviews. ***Journal of Software***, 8(11): 2682-2687. doi:10.4304/jsw.8.11.2682-2687

O'Connor, P. 2008. User-generated content and travel: A case study on Tripadvisor.com. ***Information and communication technologies in tourism 2008***: 47-58.doi:10.1007/978-3-211-77280-5_5

Pan, B., MacLaurin, T., & Crotts, J. C. 2007. Travel blogs and the implications for destination marketing. ***Journal of Travel Research***, *46*(1): 35-45.

Pang, B., & Lee, L. 2005.Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. ***Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*****:** 115-124. Association for Computational Linguistics

Pang, B., & Lee, L. 2008.Opinion Mining and Sentiment Analysis. ***Foundations and trends in information retrieval,*** 2(1-2): 1-135.

Pekar, V., & Ou, S. 2008. Discovery of subjective evaluations of product features in hotel reviews. ***Journal of Vacation Marketing***, 14(2): 145-155.doi:10.1177/1356766707087522

Pudliner, B. A. 2007. Alternative literature and tourist experience: Travel and tourist weblogs. ***Journal of Tourism and Cultural*** *Change*, *5*(1): 46-59.

Rossetti, M., Stella, F., Cao, L., & Zanker, M. 2015. Analysing User Reviews in Tourism with Topic Models. ***Information and Communication Technologies in Tourism 2015****:* 47-58**.**Springer International Publishing.

Šerić, M., Gil-Saura, I., & Ruiz-Molina, M. E. 2014. How can integrated marketing communications and advanced technology influence the creation of customer-based brand equity? Evidence from the hospitality industry. ***International Journal of Hospitality Management***, 39: 144-156. doi:10.1016/j.ijhm.2014.02.008

Shi, H. X., & Li, X. J. 2011. A sentiment analysis model for hotel reviews based on supervised learning. In Machine Learning and Cybernetics (ICMLC), ***2011 International Conference***, vol. 3: 950-954. IEEE. doi:10.1109/ICMLC.2011.6016866

Soper, D. S., & Turel, O. 2012. An n-gram analysis of Communications 2000--2010. ***Communications of the ACM***, 55(5): 81-87. doi:10.1145/2160718.2160737

Stieglitz, S., & Dang-Xuan, L. 2013. Social media and political communication: a social media analytics framework. ***Social Network Analysis and Mining***, 3(4): 1277-1291. doi:10.1007/s13278-012-0079-3

Stokes, D., & Lomax, W. 2002. Taking control of word of mouth marketing: The case of an entrepreneurial hotelier. ***Journal of Small Business and Enterprise Development***, 9(4): 349–357.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. 2011. Lexicon-based methods for sentiment analysis. ***Computational Linguistics***, 37(2): 267–307.

Turban, E., Sharda, R., Delen, D., & Efraim, T. 2011. ***Decision support and business intelligence systems***, 9th Edition. Pearson.

Turney, P. D. 2002. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. ***Proceedings of the 40th annual meeting on association for computational linguistics*****:** 417-424. Association for Computational Linguistics.

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

Xie, K. L., Zhang, Z., & Zhang, Z. 2014. The business value of online consumer reviews and management response to hotel performance. ***International Journal of Hospitality Management***, 43: 1-12.doi:10.1016/j.ijhm.2014.07.007

Wang, G., Sun, J., Ma, J., Xu, K., Gu, J. 2014. Sentiment classification: The contribution of ensemble learning, ***Decision Support Systems***, 57: 77-93.

Witten, I. H., & Frank, E. 2005. ***Data mining: Practical machine learning tools and techniques***. San Francisco: Morgan Kaufmann.

Wong, K. F., Xia, Y., Xu, R., Wu, M., & Li, W. (2008). Pattern-based opinion mining for stock market trend prediction. ***International Journal of Computer Processing of Languages***, *21*(04): 347-361.

Ye, Q., Zhang, Z., & Law, R. 2009. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. ***Expert Systems with Applications***, 36(3): 6527-6535. doi:10.1016/j.eswa.2008.07.035

Yi, X., & Allan, J. 2009. A comparative study of utilizing topic models for information retrieval. ***Advances in Information Retrieval***: 29-41. Springer Berlin Heidelberg.doi:10.1007/978-3-642-00958-7_6

Appendixes

Annex 1 - Areias do Seixo Description

"Areias do Seixo Charm Hotel & Residences" comes from a dream that the owners, Gonçalo and Marta, had of building an unique place, harmoniously integrated with its local environment and break the actual thoughts about 5 star hotels. It is a thematic luxury eco-hotel, who has already been described as an "Eco hippie techno chic" place.

This project was initially projected and built having a strong environmental awareness, which is why sustainability is one of its main values. With a unique design involving nature, the most part of the decor is made out recycled objects, offering a sustainable management of resources with a view to reducing its ecological footprint. The owner's passions and a lot of their trips around the world are present all over the hotel, where this place is not only about design and details but also about transmitting feelings and emotions, through its originality, comfort and sophistication perfectly blended with the landscape.

The place is a low density unit containing only 14 rooms, in which each room tells a different story. The décor, the materials use, the different senses create unique sensations that make guests want to return to live a different story.

The owners bet on this location, in Santa Cruz, county of Torres Vedras and district of Lisbon, especially in the West, because they feel this project deserved an improbable location away from any touristic circuit, to make the project as unique as its location.

Besides the hotel, it was also built a housing zone, with 18 houses, which works as a little resort managed by the hotel, allowing guests to take advantage of all common places and hotel services, as the pool, SPA and restaurant.

This environmental framework, also economic and social, forms a guest's segmentation which makes them prefer this place to any other. A substantial part of costumers come from Northern Europe but the intern demand is also rising. The main target is couples, as the hotel is seen as a romantic place, not only for its décor, but for its SPA and outdoor activities, as nature pathways.

Annex 2 - Initial Hotel's Director interview

What's next for Areias do Seixo how is it going to continue pushing the sustainability boundaries?

We have implemented a monitoring continuous work in sustainability. Our major concern right now is focusing on lower the consumption per person (in terms of electricity and water), an internal goal that is emphasis on resources reuse. For that, we have an even greater goal to achieve, which is, changing mentalities (of resource usage by guests mainly), without of course hurting susceptibilities of someone who is paying 300€ for a room.

Gonçalo reinforces the balance between sustainability and comfort. And that promoting this kind of behavior in today's society is a hard thing to do, considering that people are used to some privileges nowadays that some years ago they had not. The main question here is that people can take advantage of this privilege but in a more balanced way, balancing comfort with sustainable methods. They are trying harder to capture this mindset in their guests, reinforcing that the majority of them already has that "green mindset", mainly the foreigners. But Gonçalo also considers Portugal is making big progresses in this area.

What is in terms of Marketing its main means?

The type of marketing implemented in this project is the marketing named "word of mouth". Meaning that the guests itself are going to promote the place by themselves, nationally and internationally, in opinion sites, blogs, friends talks. "We have made our guests our public relations", states Gonçalo.

That are several guests who say "This is so special that I don't feel like I want to tell anyone" on which Gonçalo answer: "Don't tell everyone, tell the right people". This reinforces the singularity and exclusiveness of this place.

Do you consider environmental responsible behavior adopted in this project have direct impact in their financial performance?

Gonçalo reinforces the substantial initial investment in "green technologies" (as solar panels, geothermic installation, among other technologies), but that in a middle-long term this investment is already paid off. Proving that by adopting a sustainable strategy it is necessary a substantial initial investment that would not pay right away but long term the hotel will have their income and its financial performance increasing for sure.

He also says that after its opening in 2010, after 5 years, the initial investment is paid off and now is only taking advantage of its benefits. Which reinforces resource's efficiency and the differentiation of the project, along with the right "pricing" established.

What are, if you have to name, the main competitors of Areias do Seixo?

Gonçalo states which their main influencers outside of Portugal were: The Brando and Six Senses. And in Portugal he says that there are few concepts close to what this project is, but the ones which are closer to this concept and can be considered competitors are Rio Prado in Óbidos, and Santa Marta Inspira in Lisbon.

How is the process of parameterization/evaluation of the performance of the hotel?

We have some monitoring follow up grids, distinguishing the strong from the weakest points. In terms of reception, restaurant, housekeeping, maintenance, SPA and Others, that we are able to obtain through our guest's opinions in multiple channels, namely on Trip Advisor, Guest Book, Follow-up emails, evaluation websites and from direct emails. The main goal of this grid is to get the Quality certification, considering that questionnaires to guests are not allowed in the Hotel, once it is not our policy, says Gonçalo.

What are the main reasons for people to choose this hotel?

What have raised interest on people to visit this place was in first place to watch a picture of it, probably in some international/national magazine, on the Internet or in a Blog. Mainly they saw on a place that as the strength of an opinion maker.

Figure 6 - Process of choosing Areias do Seixo as an Hospitality destination according to Hotel' Director

Do you consider guests choose a hotel based on its environmental friendly conscious?

Gonçalo states that is not only sustainability that makes guest to come here, above that is the architecture and decoration factor. When people realize that this is a sustainable project it only gives them more will to come. He also states that the majority of its guest likes to be engaged on environmental activities, taking into consideration that big part of them are people who came from different cultures and have this sustainable factor more intrinsic.

What do you consider the main reasons for not adopting environmental behavior when travelling, or don't go to Areias do Seixo?

Gonçalo argues that the price is clearly the main reason for people to don't stay in this type of places. He also considers that is a high value for money. And people in general consider it expensive.

Do you think eco-labelling is just used as a Marketing tool or acknowledgements of Areias do Seixo practices?

Gonçalo states with confidence that guest's would consider that those eco-labels and certification of the place are all acknowledgements of Areias do Seixo practices.

Some of the activities/experiences offered to guest includes the participation in organic agriculture, namely in Areias do Seixo garden. Where the majority of products, are then, used in the preparation of restaurant meals. Another activity is based on picking mussels and walk on the dunes; another incentives the guests to reuse organic waste and "think eco" as guest participate in Hotel' Eco circuit.

How do you describe your typical client?

The typical client is considered an urban person, from big capitals, who wants to run away of quotidian inquietude, he/she seeking a retreat space which connects with nature and what are the essentials of life.