Management Letters / Cuadernos de Gestión 25/1 (2025) 71-88

FESIDE
Fundación Emilio Soldevilla para la Investigación y Desarrollo de la Economía de la Empresa

enpresa institutua
Instituto de Economía Aplicada a la Empresa

Management Letters / Cuadernos de Gestión

journal homepage: https://ojs.ehu.eus/index.php/CG

ISSN: 1131-6837 / e-ISSN: 1988-2157

Improving guest satisfaction by identifying hotel service micro-elements failures through Deep Learning of online reviews

Mejora de la satisfacción del cliente mediante la identificación de fallos en los microelementos del servicio hotelero mediante el Deep Learning de las reseñas online

Sergey Kazakov[*], Pedro Cuesta-Valiño[a], Vera Butkouskaya[b], Daniel Muravsky[c]

[a] *University of Alcalá. Faculty of Economic, Business and Tourism Sciences. Plaza de la Victoria, 2 28802 Alcalá de Henares (Madrid), Spain – pedro.cuesta@uah.es – https://orcid.org/0000-0001-9521-333X*

[b] *Graduate School of Business, HSE University. Tecnocampus, Universitad Pompeu Fabra. Carrer d'Ernest Lluch, 32, 08302 Mataró, Barcelona, Spain – vbutkouskaya@tecnocampus.cat – https://orcid.org/0000-0002-6963-3872*

[c] *University of West Scotland, School of Business and Creative Industries. Import Building, 2 Clove Cres, London E14 2BE, United Kingdom – Prof.muravsky@gmail.com – https://orcid.org/0000-0001-8401-1476*

* ***Corresponding author:*** *University of Alcalá. Faculty of Economic, Business and Tourism Sciences. Plaza de la Victoria, 2 28802 Alcalá de Henares (Madrid), Spain – sergey.kazakov@uah.es – https://orcid.org/0000-0001-5532-5791*

ARTICLE INFO

Received 21 February 2024,
Accepted 18 Aug 2024

Available online 21 March 2025

DOI: 10.5295/cdg.242191sk

JEL: M31

ABSTRACT

This study thoroughly examines often-overlooked micro-service elements within the broader spectrum of hotel services, aiming to improve hospitality services and ensure guest satisfaction. To achieve this, this research developed a methodological framework, integrating (a) the VADER text sentiment analysis framework, (b) a robust logistic regression procedure to pinpoint specific hotel service components culprit for guest frustration, and (c) the application of semantic network analysis to yield guest insights contextualised within the realm of underperforming hotel service micro-elements.

Research findings highlight fifty specific service micro-elements identified as triggers of negative sentiment and subsequent degrees of diminished guest satisfaction. Furthermore, this study zooms into the top ten underperforming service micro-elements by employing semantic network analysis to uncover the roots of typical guest frustrations with their hotel experiences. Though identified within hotel reviews, certain service malfunctions have relevance within the broader domain of destination management.

The outcomes of this study suggest a valuable resource for managers in detecting and rectifying inadequately performing hotel service micro-elements, which are pivotal for elevating guest satisfaction within their respective hotel properties. Additionally, the findings provide impetus for hotel and destination managers to implement tailored strategies to increase guest satisfaction across hotels and destinations.

Keywords: hotel service elements, online reviews, natural language processing, big data, tourist satisfaction policy, eWOM.

This article is distributed under the terms of the *Creative Commons Atribution 4.0 Internacional* License

R E S U M E N

Este estudio examina en profundidad los elementos de microservicios a menudo pasados por alto dentro del amplio espectro de servicios hoteleros, con el objetivo de mejorar la hospitalidad y garantizar una mayor satisfacción de los huéspedes. Para lograrlo, se desarrolló un marco metodológico que integra (a) el análisis de sentimiento de texto VADER, (b) un procedimiento robusto de regresión logística para identificar los componentes específicos del servicio hotelero que causan frustración a los huéspedes, y (c) el análisis de redes semánticas para generar información matizada sobre los huéspedes, contextualizada dentro del ámbito de los microelementos de servicio hotelero de bajo rendimiento.

Los resultados de la investigación destacan cincuenta microelementos de servicio específicos que desencadenan sentimientos negativos y una disminución subsecuente en la satisfacción de los huéspedes. Además, este estudio se enfoca en los diez microelementos de servicio de menor rendimiento, utilizando el análisis de redes semánticas para descubrir las causas principales de las frustraciones comunes de los huéspedes con sus experiencias hoteleras. Algunos fallos en el servicio, aunque se identifican en las reseñas de hoteles, son relevantes también en el ámbito más amplio de la gestión de destinos.

Los hallazgos de este estudio sugieren un recurso valioso para los gerentes en la detección y corrección de microelementos de servicio hotelero que funcionan de manera inadecuada, fundamentales para elevar la satisfacción de los huéspedes en sus respectivas propiedades hoteleras. Además, los resultados incentivan a los gerentes de hoteles y destinos a implementar estrategias personalizadas destinadas a mejorar la satisfacción de los huéspedes en todos los hoteles y destinos.

Palabras clave: elementos de servicio hotelero, reseñas online, procesamiento del lenguaje natural, big data, política de satisfacción del turista, eWOM.

1. INTRODUCTION

Online reviews have become a vital channel for hotel guests to share their travel experiences (Xie *et al.*, 2014; Casalo *et al.*, 2015). These reviews are crucial in the customer journey, often representing the final stage in tourism and hospitality settings (Chatterjee, 2020). They provide valuable content for prospective hotel customers (Repovienė & Pažėraitė, 2023) and represent electronic word-of-mouth (eWOM), a digital extension of the traditional word-of-mouth (WOM) concept in the domain of conventional marketing (Butkouskaya *et al.*, 2020; Le *et al.*, 2023). As such, online reviews significantly influence hotel booking decisions, presenting immense opportunities for hospitality researchers to explore various facets of guest behaviour (Yang *et al.*, 2020). Consequently, advancements in ICT, easy access to online reviews as research data, and machine learning technologies have fueled a proliferating stream of big data-based research in tourism and hospitality (Cuesta-Valiño *et al.*, 2020).

In their reviews, hotel guests frequently comment on the services they encounter, including hotel rooms and facilities, personnel, services, location, and food (Nie *et al.*, 2020; Zarezadeh *et al.*, 2022). Simultaneously, hotel guests commonly evaluate these service macro-factors and focus on specific service micro-elements within the above-noted categories. For instance, when assessing a hotel room, guests often highlight aspects like cleanliness, bed quality, Wi-Fi, and bathroom amenities (Luo *et al.*, 2021).

Given their impact on overall guest evaluations, hotel service micro-elements are meaningful and thus require more consideration in hospitality marketing literature. We define hotel service micro-elements as specific and detailed facets of the hospitality servicescape that collectively shape the overall guest experience. Although previous research has examined micro-elements' effects on guest satisfaction (Hu *et al.*, 2020), it often used the term' attributes' without distinguishing between different levels of service elements.

In more detail, recent studies have operated with service macro factors, referred to as 'attributes', to all elements of the hotel's servicescape without making a distinction. For instance, prior literature has positioned such service elements as 'bed', 'bathroom', 'internet' and 'room' or 'drink', 'lobby' and 'bar' on the same level. However, there is an apparent hierarchy where 'room' is a higher order element comprising 'bed', 'bathroom', and 'internet'. At the same time, 'drink' belongs to 'bar', which sequentially fits in the 'lobby', a larger-scale service attribute. Without making a multi-level taxonomy between the sensual dimensions of the hotel service elements, it becomes challenging for hotel managers to determine an accurate, sometimes hidden reason for guests' disappointment with the hospitality service. Moreover, as the perception of the higher-order service elements (room, for instance) denotes a complex combination and likely implies a regressed sum of the lower-level elements (bed, air conditioning, furniture, etc.), considering guests' hotel evaluation should be multifaceted and non-linear given from this perspective also.

Despite some amount of extant research on guest satisfaction using advanced techniques of text mining and natural language processing (NLP) applied to online guest reviews as a data source(e.g., Hu & Yang, 2021; Shin *et al.*, 2021), the above-noted limitations generate a significant research gap that is awaiting academic attention. Apart from confusion due to the noted service elements failures examination without considering the elements' hierarchy, prior studies have also often been limited to specific destinations, such as China. Consequently, the literature must still address the research findings' generalisation issues. Moreover, studies in the field of hospitality customer behaviour should also provide better in-depth and more concrete insights into the reasons for negative sentiments discovered in guest reviews.

Addressing negative customer feedback is crucial, as negatively tinted reviews significantly impact hotel perceptions of prospective customers (Hu *et al.*, 2020). Grounded in the accumulated research, its above-noted accomplishments and gaps determined in the extant body of literature, this study poses the following research question:

RQ: *According to online reviews, which specific hotel service micro-elements generate the most negative sentiment, lead to poor customer experiences, and disrupt guest satisfaction?*

To address the posed RQ, this exploratory study aims to (1) use deep learning and text mining to identify hotel service micro-elements that contribute to negative guest sentiment in online reviews across eleven popular tourist destinations and (2) determine the failing service micro-elements and explore the context of these poorly delivered services to understand the underlying causes of guest dissatisfaction. By addressing the posited RQ, this study makes two significant contributions: First, in the realm of hospitality marketing literature, it determines the top failing hotel service micro-elements and their contexts overlooked by prior studies. Second, this research provides hotel managers with information on specific service flaws requiring their attention. Also, this study bestows detailed practical recommendations to improve service quality and guest experiences.

The remaining manuscript is organised as follows. First, we review the literature that underpins our research rationale and supports the posed RQ. Next, this paper presents the research methodology, explicating the developed approach to data collection and analysis procedures. Then, the narrative presents the findings and discusses the implications for hospitality marketing theory and hotel management practices. Finally, we outline research limitations and propose future research directions, concluding with a summary of our key findings.

2. LITERATURE REVIEW

2.1. Online guest reviews and hotel service failures

As previously noted, several researchers indicated that online guest reviews are a practical data source for collecting guests' insights to evaluate a customer's perception of the hotel service quality (Berezina *et al.*, 2016; Song *et al.*, 2022; Zarezadeh *et al.*, 2022). Hotel guests share their experiences with other users by posting reviews to social media, portals of online travel agencies (OTAs) or review-specialised platforms (Khan *et al.*, 2022; Vieira *et al.*, 2023). Online reviews denote a piece of user-generated content (UGC) commonly comprising both qualitative (review text) and quantitative (guest-assigned hotel rating score) data (Mariani *et al.*, 2019). As a valuable source of public data, online reviews help enact analytical techniques such as social media listening, big data analytics and text mining (Hu & Yang, 2021).

Negative reviews posted by hotel guests are of particular research interest because they allow hotel managers to identify service failures and, thus, find opportunities to improve services delivered by their property (Son *et al.*, 2022). Researchers argue that hotel service failure refers to situations when the hotel services do not comply with the guests' expectations, leading to a significant deviation from the expected service standards (Sann *et al.*, 2021). An extant body of amassed research has used text mining and determined several factors guests associate with hotel service failures (e.g., Huang *et al.*, 2022; Nie *et al.*, 2020; Ying *et al.*, 2020).

In this vein, researchers first pointed to room cleanliness as the critical micro-element predisposing negative low guest satisfaction (Park *et al.*, 2019). Prior research has also noted hotel facilities in online reviews concerning service failures (Ying *et al.*, 2020). Furthermore, according to the literature, guests indicate customer service provided by hospitality personnel in their negative reviews as a factor disrupting their satisfaction with the hotel (Nie *et al.*, 2020).

2.2. Hotel service micro-elements

Various marketing researchers have explored how customers perceive service elements across receiving customer experiences and satisfaction with service products (Bueno *et al.*, 2019; Roy, 2018). In hospitality literature, amassed research has determined drivers influencing guest satisfaction with hotel services (Lee *et al.*, 2020). In this vein, as previously noted, several researchers pointed to the hotel room as the critical factor predisposing negative sentiment and, thus, low guest satisfaction (Padma & Ahn, 2020; Park *et al.*, 2019). Next, Latinopoulos (2020) attributed hotel location and exterior as a highly influential element of the hotel service. Furthermore, guests indicate food quality as a factor sculpting their satisfaction with the hotel (Philips *et al.*, 2017; Zarezadeh *et al.*, 2022).

Simultaneously with the established research focusing on the overarching service factors relevant to guest satisfaction, a growing body of literature delves into the finer or micro-elements of hotel service elements. In this regard, Nie *et al.* (2020) revealed sleep quality as a hotel room's micro-element, which is significant to the guests. Later, Hu *et al.* (2021) echoed these findings but unveiled an extended range of hotel micro-services influencing guest satisfaction, including room rate, lobby bar, reception staff, breakfast, and wi-fi signal quality. Simultaneously, Luo *et al.* (2021) noted service micro-elements of the wi-fi signal, air conditioning, bed, noise, towels, hairdryers, and slippers priorly examined as bigger 'room' or 'room facilities' factor by other researchers (Alnawas & Hemsley-Brown, 2018). More recently, Song *et al.* (2022) used the Latent Dirichlet Allocation (LDA) algorithm to extract topics from more than 50 000 guest reviews. In that study, LDA extracted five topics that matter to guest satisfaction, including Service, Room, Cleanliness, Location, and Value (Song *et al.*, 2022). Prior researchers have widely implemented sentiment analysis, discussed below, as a text mining technique to obtain these findings. By employing sentiment analysis, they extracted negative guest reviews first and, by applying additional scrutiny, found the above-noted hotel service failure factors that commonly reduce guest satisfaction with the hotel (Huang *et al.*, 2022).

2.3. Sentiment analysis in hospitality research

Sentiment analysis is a text mining procedure researchers use to narrow down a rich palette of human emotions, including happiness, pleasure, apathy, annoyance, rage, etc., into three distinct categories of negative, neutral or positive polarity (Kirilenko *et al.*, 2018; Luo *et al.*, 2021). Online guest reviews denote a big data solution for hospitality researchers that allows them to gauge review sentiment more accurately and depict review polarity with a precise numeric value. Online guest reviews benchmark a substantial move forward in hospitality analytics and gradually replace conventional methods of data collection, e.g. surveys, required for obtaining customer sentiment data (Hu *et al.*, 2020). By applying the text mining technique and fundamental sentiment analysis procedures, hospitality researchers look for specific words or complete phrases in customer reviews which determine a guest's positive or negative sentiment. Accordingly, several studies have exhibited visually impressive word cloud models grounded in content frequency analysis of reviews (Hu & Yang, 2021; Shin *et al.*, 2021).

Sentiment analysis is a part of the Natural Language Processing (NLP) algorithms family, which is a ramification of the broader machine learning procedures constellation. Thanks to its versatile benefits, lexicon-based sentiment analysis has received recognition as a robust methodology in academia. It has set the momentum for an emerging tourism and hospitality research stream. In this vein, Yadav and Roychoudhury (2019) applied sentiment analysis to investigate hotel attributes that travellers perceive as essential when planning their trip in different travel modes (leisure, business, single, couple, family, etc.). Similarly to these findings, a study by Berezina *et al.* (2016) revealed common categories used in positive and negative reviews by applying text mining techniques to analyse online customer reviews. Moreover, several studies utilised sentiment analysis to develop predictive models for computing numerical rating scores missing in the review dataset (Geetha *et al.*, 2017; Kim & Im, 2018).

Next, sentiment analysis demonstrated its cogent capabilities to determine the linkage between the importance, performance, and customer satisfaction of hotel service attributes, according to Hu *et al.* (2020). Furthermore, in this regard, Luo *et al.* (2021) invoked sentiment analysis in examining economy hotels in China and determined hotel service elements that negatively influence customer experience. Another notable study by Nie *et al.* (2020) suggested a unique approach to building a hotel recommendation system grounded in blending multiple criteria decision-making (MCDM), sentiment analysis, and latent Dirichlet allocation (LDA). These noted studies have demonstrated that sentiment analysis proved a robust and reliable text mining and analytical tool in hospitality research to examine and determine the antecedents of hotel guest satisfaction.

3. RESEARCH METHODOLOGY

To address the posed RQ and meet this exploratory study's aims, the developed research methodology comprised five sequential data collection and engineering phases, applying a series of data analysis procedures and visualising the results. Figure 1 depicts the developed methodology and the logic we followed in this research.

3.1. Sampling procedure

This research used a simple random sampling (SRS) procedure to establish a research setting for this study. It is assumed that SRS cannot tackle notorious generalisation issues because destinations, hotel service attributes, and, thus, guest experiences may vary (Malina *et al.*, 2011). Nevertheless, such an approach complies with the accuracy of the empirical study design in a single setting to get generalisable customer insights into the hotel services belonging to a particular tourist destination (Mariani *et al.*, 2019).

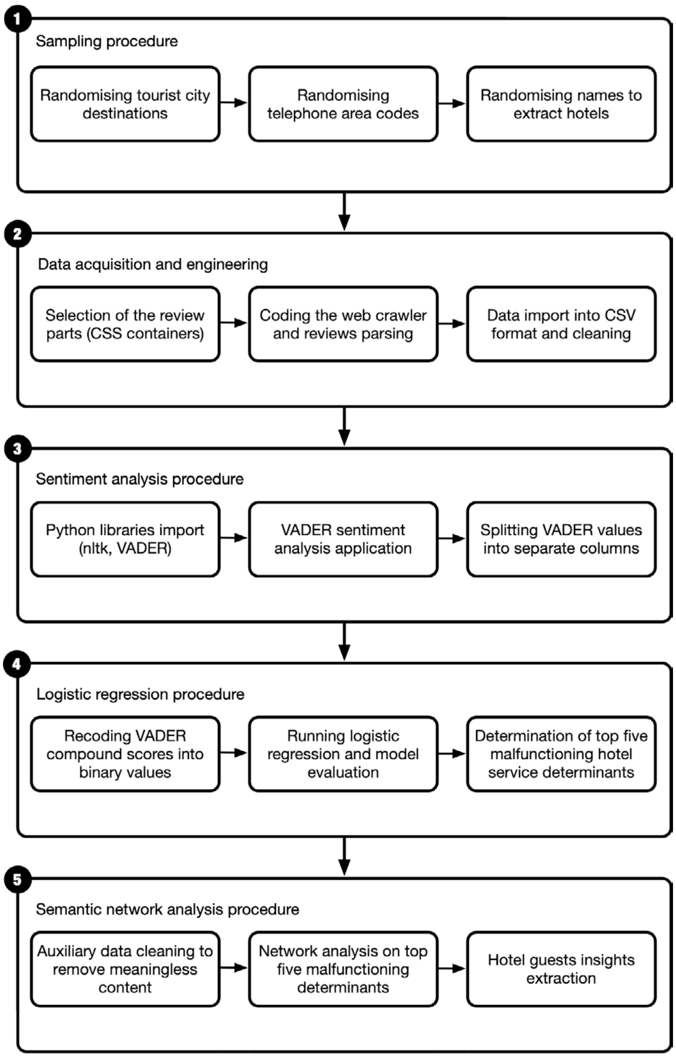

Figure 1
Research methodology scheme
Source: Own elaboration.

The present study employed a four-stage SRS procedure. For this purpose, we first applied an online random digit generator (https://random.org/) to pick the number between 5 and 20 out of the 100 most visited tourist city destinations in 2022. One hundred cities were considered because the tourism analysis literature commonly employs this particular number of top performers in calculating the city destinations index (Popova, 2023). Also, it was assumed that most visited cities were a good source of sufficient guest review availability required for analysis. The generator had returned 11, so this research collected data from eleven destinations to be defined further by SRS. As Baltes and Ralph (2022) suggested for SRS procedures, we used a random draw from the city area phone codes of 100 popular destinations to determine 11 cities for data collection (Table 1). Next, the random digit generator picked 4-star hotels from the 2-5 stars option. We then opted for booking.com as a custom data source platform commonly utilised in hospitality research (Sann *et al.*, 2021).

3.2. Data collection

This research employed Python *requests, beautifulsoup*, and *lxml* libraries to scrape and parse the hotel guest reviews and build a review text corpus for further analysis. We configured the crawler to scrape the significant parts of reviews to originate the variables and their respective values required according to the developed research design. The targeted data included hotel name, country of a reviewing guest, hotel rating score given by the reviewer, and review text itself — its 'positive' and 'negative' parts, as booking.com splits the guest review form into these two categories. The web crawler was programmed to search and scrap reviews relevant to four-star hotels in all written languages in eleven destinations. Next, according to the suggestions of Mariani *et al.* (2019), the crawler was set to skip blank or incomplete reviews where the text length was less than 15 words. The year 2022 was a period to be considered by the web crawler. After its run, the crawler scraped N=109715 hotel guest reviews to a tabulated dataset from 11 city destinations determined by the SRS procedure (Table 1).

Table 1
City destinations and number of scrapped reviews

##	Destination	Country	n of reviews
1	Bordeaux	France	2270
2	Dubai	OAE	10550
3	Hawaii	USA	10352
4	Khurgada	Egypt	12902
5	Las Vegas	USA	11429
6	London	UK	11556
7	Malaga	Spain	10156
8	Munich	Germany	3345
9	New York	USA	17978
10	Prague	Czech Rep.	8880
11	Qatar	Qatar	10297
		Total n of reviews:	**109715**

Source: Own elaboration.

As the design of this study required the text parts of the reviews solely to complete sentiment analysis, the country of the reviewer and hotel rating score variables were discarded from the current analysis for the needs of future studies. The obtained data required auxiliary data engineering to receive a complete dataset ready for analytical procedures. In this vein, as the research design required the sentiment of the entire guest review to be gauged, we first concatenated columns in the retrieved dataset containing available text parts into a single variable representing the full guest review (Figure 2).

```
В [7]: data.head(10)
Out[7]:
```

	Hotel	review	country	score
0	soho_boutique_las_vegas_malaga	An enjoyable stay with great staff, breakfast ...	Spain	8.0
1	soho_boutique_las_vegas_malaga	Very good The room I had was a little dated bu...	Spain	8.0
2	soho_boutique_las_vegas_malaga	Considering we are in a pandemic, we had a lov...	UK	9.0
3	soho_boutique_las_vegas_malaga	Superior As it was our daughters birthday they...	Ireland	9.0
4	soho_boutique_las_vegas_malaga	Fabulous short break sun sea culture very beau...	Ireland	9.0
5	soho_boutique_las_vegas_malaga	Great location and nice hotel. Great location ...	UK	9.0
6	soho_boutique_las_vegas_malaga	This hotel is really handy for the city but al...	UK	9.0
7	soho_boutique_las_vegas_malaga	Very good Location to the beach There was a lo...	UK	8.0
8	soho_boutique_las_vegas_malaga	Nice friendly place in pleasant location. Very...	UK	8.0
9	soho_boutique_las_vegas_malaga	Very good Nice little hotel. Good location, cl...	UK	8.0

Figure 2
Parsed data set fragment
(dataset with concatenated review before removal of unnecessary data columns)
Source: Own elaboration.

The mined guest reviews were written in different national languages. Whereas Python natural language processing (NLP) frameworks demonstrate their best capabilities when applied to texts in English, it was essential to translate review texts into this particular language. Python NLP VADER sentiment analysis library, applied in this research and which we depict below in this section, provides such translation options. VADER executes text translation to English by utilising a pre-trained Google Cloud Translate API base with a high level of translation accuracy, according to literature Wang, L., & Kirilenko, A. P. (2021). On top of that, we additionally employed two other online translators, IBM Watson Language Translator and Yandex.Translate to ensure a correct translation on ten randomly selected reviews for every language found in the dataset. This procedure helped verify no loss of meaning, as all tested texts had the same denotation in English.

3.3. *Text data preprocessing*

Almost all the machine learning tasks relevant to NLP require additional procedures to engineer collected text data for subsequent analysis. These procedures, at first, help to split text data into smaller chunks, referred to as tokens, and then apply auxiliary techniques to ensure the precision of the research results (Perkins, 2014). This research employed ordinary data preprocessing procedures relevant to NLP, as Alam and Yao (2019) suggested. These essential NLP procedures can be run by the *nltk* (referred to as *Natural Language Toolkit*) Python framework, recognised as an industry-standard solution for text data mining procedures.

We commenced the application of NLP preprocessing techniques with text tokenisation (also known as *lexing*) to convert words into measurement units, or tokens, by removing punctuation and whitespaces, making them ready for analytical manipulations. For this purpose, we favoured the UPPipe tokeniser, which is noted as a versatile and reliable tool for text tokenisation according to the extant body of literature (Straka & Straková, 2017). Furthermore, after tokenisation, we sequentially utilised three procedures, namely, (a) text transformation by eliminating letter accents found in some words; (b) text normalisation where we ran stemming geared by the UDPipe lemmatiser (Straka *et al.*, 2016); and (c) text filtering to discard regular expressions and stop-words from the text. These procedures were essential to reduce bias and attain more precise results for analysing the reviews' text data.

3.4. *Data analysis procedures*

This study utilised three sequential data analytical techniques comprising sentiment analysis, logistic regression and semantic network analysis (Figure 1). Sentiment analysis was significant for this research to sort out and zoom in on the negative reviews that would serve as a source to reveal the original guests' insights on their hotel experiences. These insights were essential to discern service elements that routinely caused negative customer experiences in the sampled hotels.

For sentiment analysis, this study has applied VADER (Valence Aware Dictionary for Sentiment Reasoning), a lexicon-based solution in NLP capable of gauging polarity and its valence simultaneously. Thanks to these advantages, the researcher receives more precise sentiment score values. This is plausible because the VADER framework is not limited to relying solely on a lexicon while making its computations; it is also a rule-based application. In summary, this means that VADER reckons sentiment valence by the context of a specific word while computing its sentiment and considering capitalisation and even *emojis* if they accompany words or are standalone in the analysed text. These VADER's capabilities make it a more reliable approach to receiving precise sentiment values than the other frameworks (Hutto & Gilbert, 2014).

Next, the subsequent data analysis stage applied the logistic regression technique. It was necessary to reveal the most impactful hotel service elements that precipitate low guest satisfaction. As logistic regression requires a binary dependent variable, we normalised VADER's continuous compound sentiment score to 0 and 1 to meet this requirement. A value of «0» represented a negative guest review sentiment, while «1» denoted a positive sentiment in the updated dataset. Although the logistic regression helps highlight the hotel service micro-elements causing bad guest experiences, it still provides a paucity of information for generating proper insights into service failures.

Conversely, the uncovered context of the found service micro-elements is a practical means to obtain more information on the antecedents of guest frustrations and discern the details of hotel service failures. For this reason, this research employed semantic network analysis (Oh & Kim, 2020) to shed more light on the context of malfunctioning hotel service micro-elements. By aiming to obtain a more precise comprehension of such context, we implemented semantic network analysis solely on the reviews with negative sentiment scores following a suggestion from prior literature (Bachleda & Berrada-Fathi, 2016; Israeli *et al.*, 2019). Such reviews were sorted from the initial dataset and put into a subsample of n = 23135. The next section of this paper depicts the results attained after applying semantic network analysis and the rest of the above-explicated data analysis procedures.

4. RESULTS

4.1. Guest review sentiment analysis

As noted earlier, this study has employed the VADER framework to implement sentiment analysis. VADER algorithm analyses text and returns four sentiment scores: negative, neutral, positive, and compound. These scores have values in the range of —1 to 1. VADER computes the compound sentiment value by normalising the sum of the other three score values (Hutto & Gilbert, 2014). The outcomes of the VADER procedure application for sentiment analysis are depicted in Figure 3.

```
B [20]: data.head()
Out[20]:
```

	Hotel	review	country	score	vader_scores	vader_compound	vader_pos	vader_neg	vader_neu
0	soho_boutique_las_vegas_malaga	An enjoyable stay with great staff, breakfast ...	Spain	8.0	{'neg': 0.057, 'neu': 0.701, 'pos': 0.242, 'co...	0.9573	0.242	0.057	0.701
1	soho_boutique_las_vegas_malaga	Very good The room I had was a little dated bu...	Spain	8.0	{'neg': 0.0, 'neu': 0.608, 'pos': 0.392, 'comp...	0.9919	0.392	0.000	0.608
2	soho_boutique_las_vegas_malaga	Considering we are in a pandemic, we had a lov...	UK	9.0	{'neg': 0.05, 'neu': 0.759, 'pos': 0.191, 'com...	0.9835	0.191	0.050	0.759
3	soho_boutique_las_vegas_malaga	Superior As it was our daughters birthday they...	Ireland	9.0	{'neg': 0.0, 'neu': 0.685, 'pos': 0.315, 'comp...	0.8881	0.315	0.000	0.685
4	soho_boutique_las_vegas_malaga	Fabulous short break sun sea culture very beau...	Ireland	9.0	{'neg': 0.035, 'neu': 0.605, 'pos': 0.36, 'com...	0.9344	0.360	0.035	0.605

Figure 3

Reviews dataset excerpt with VADER sentiment analysis output values.

Note: This figure represents the first five rows of the entire dataset, demonstrating the output of running the VADER procedure.

Legend: 'Hotel' – hotel to which guest review was posted; 'review' – guest review text displayed partially due to the Python client limitations; 'country' – country of guest nationality; 'score' – hotel rating score set by guests; 'vader_scores' – Python variable of dictionary type {'key': 'value'} containing computed sentiment values for each guest review; 'vader_compound' - a normalised sum of VADER's positive, neutral and negative sentiment score value; 'vader_pos' - VADER's positive sentiment score value; 'vader_neg' - VADER's negative sentiment score value; 'vader_neu' - VADER's neutral sentiment score value.

Source: Own elaboration.

As indicated previously in the methodology section, obtaining sentiment score values facilitated a generation of the dependable binary variable, making it possible to implement logistic regression.

4.2. Logistic regression

The applied logistic regression was set to Ridge (L2) regularisation type, cost strength C=1, sampling cross-validation with 10 folds, and balance class distribution to reveal positive coefficient values for the hotel service elements that imply a reason for a negative review sentiment. The resulting logistic regression model unveiled adequate model evaluation statistics (Table 2).

Table 2

Logistic regression model evaluation

Model	Precision	Recall*	AUC**	CA***	F1****
Logistic Regression	0.928	0.960	0.936	0.909	0.943

Note: * — in the domain of Error type I, this value denotes the proportion of true positive values among all positive observations of the dataset;** — represents the square area below the prediction ROC (Receiver Operating Curve, Figure 4); *** — implies classification accuracy, e.g., the share of instances that were adequately classified; **** — denotes a mean of precision and recall weighted harmonically $\left(F1_{score} = 2 \cdot \frac{(Precision \cdot Recall)}{(Precision + Recall)}\right)$.

Source: Own elaboration.

ROC analysis is a cogent tool utilised to verify the accuracy of the developed model (Fawcett, 2006). It compares the model's false positive (FP) rate or specificity with the maximum probability that targets 1 while the actual value = 0 with the model's true positive (TP) sensitivity where probability targets 1 while true value = 1. Figure 4 visualises the model evaluation results using the ROC analysis. According to Fawcett (2006), the model accuracy is evident when the curve is above the dashed line representing the non-discriminatory test and near the graph's left and top borders. The ROC curve complies with this requirement, providing further evidence of the high degree of the developed model accuracy (Figure 4).

Logistic regression returned 281 tokens in the dataset as independent variables. Many of these variables needed a meaningful sense and thus were not valuable for further data analysis. To discard them, we implemented an NLP procedure known as POS-tagging to extract nouns from the variable list. As a part of speech (POS), nouns commonly represent the hotel service elements in guest reviews (Geetha *et al.*, 2017). We extracted tokens relevant to the noun POS and received 50 variables solely pursuant to the hotel service micro-elements that generate guest resentment with the hotel (Annex I).

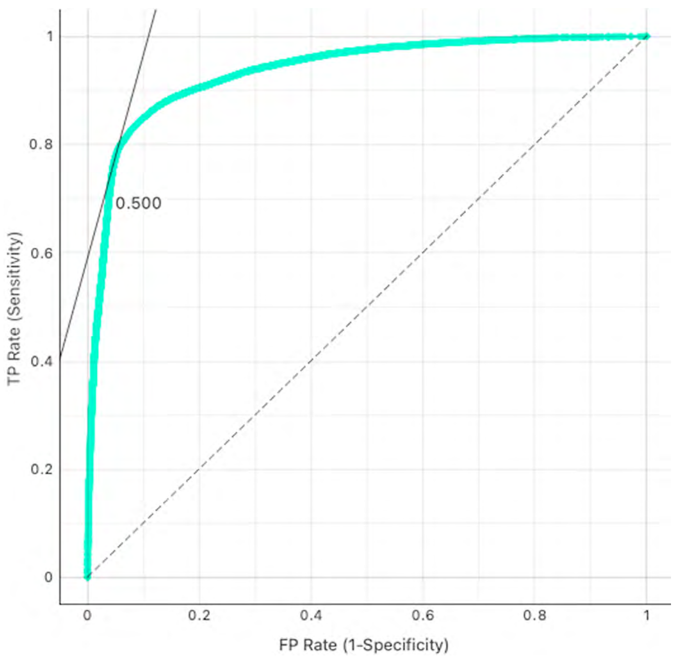

Figure 4
Logistic regression evaluation with the ROC analysis
Source: Own elaboration.

Finally, we highlighted the top ten independent variables representing failing service micro-elements to scrutinise their context and analyse the reasons for their origination (Table 3).

Table 3
Top ten hotel service elements precipitating negative review sentiment (logistic regression coefficient (β) values)

Target variable value = 0 (negative sentiment score)			
Regression intercept	**0.068**	**negative reviews count:* n = 23135**	**TF-IDF scoring μ:****
Failing service micro-elements	**β value**		
(air) conditioning	0.365	458	0.003
carpet	0.358	767	0.003
gym	0.356	162	0.001
kettle	0.331	373	0.002
table	0.307	300	0.001
luggage	0.300	902	0.004
pay	0.295	352	0.002
internet	0.268	306	0.002
pictures	0.254	399	0.002
tv	0.228	1019	0.003

* — n of guest reviews containing the service element;

** — denotes a mean of the token relevant to a hotel service micro-element computed with TF-IDF (*Term Frequency-Inverse Document Frequency*) metric that is a statistical method used to determine the significance of a word concerning a document within a set of documents.

Source: own elaboration

4.3. Semantic network analysis

At the final stage of data analytical procedures application, this research applied network analysis for the top ten poorly operated hotel service elements retrieved from logistic regression (Table 3). Before running semantic network analysis, executing an NLP technique of Concordance to build a context around the token representing a specific service microelement is essential. Concordance requires a setting of N-gram range, namely the number of tokens surrounding an analysed service microelement in every guest review where it can be found. This study employed an N-gram range of 6, meaning three tokens before and three tokens after the service microelement, to generate ample context around it to be analysed further.

The in-depth scrutiny of the semantic network analysis applied to the principal dysfunctional hotel service micro-elements reveals several findings. At first, the logistic regression application indicated that (air) conditioning (β = 0.365) is the top service microelement which deteriorates guest satisfaction. The network map (Annex II, Figure 1) divulges the details accompanying guest frustrations, including no air conditioning availability in hotel rooms and dirty, noisy, broken or improperly working appliances. Second, hotel and hallway room carpets (β = 0.358) received many guest complaints, according to our findings (Annex II, Figure 2). Guests pointed to the old, stained, dirty, worn, smelly carpeting in their hotel reviews. Third, hotel guests were dissatisfied with the gym facilities (β = 0.356). In this line, small, dirty, poorly equipped facilities and the swimming pool are reasons for guests' irritation (Annex II, Figure 3). Next, the kettle was ranked the fourth microelement contributing to the negative perception of the hotel services (β = 0.331). Semantic maps point to the kettle's unavailability in the room, faulty or broken appliances, and limited or no availability of cups, mugs, and tea sachets (Annex II, Figure 4). A hotel room furniture piece such as a table was the fifth top complaint noted in hotel reviews (β = 0.307). In this regard, guests highlighted the absence of tables or chairs and mentioned dirt found on this furniture item (Annex II, Figure 5).

Furthermore, the outcomes of this study point to luggage as the sixth top-ranked microelement, predisposing guest disappointment toward hotel services (β = 0.300). Online reviews with a negative sentiment indicate that guests may experience difficulties with luggage storage service at the reception and little help from the staff to carry heavy luggage to the hotel room (Annex II, Figure 6). Concerning payment (β = 0.295), guests' dissatisfaction arise from extra payments, need for money deposits, high pricing of the supplementary services and incorrectly charged bank cards (Annex II, Figure 7). Payment is followed by internet (β = 0.268) which generates a reduction in guest satisfaction because of wi-fi unavailability or poor, slow, and inadequate connection in the hotel room (Annex II, Figure 8). Interestingly, guests pointed to one item which is not directly linked to the hotel product consumption but pertains to the initial stage of the customer journey in hospitality. This item is relevant to pictures (β = 0.254) that hotels employ in their online marketing. Concerning pictures, hotel guests massively noted a discrepancy between the online hotel photos and the actual views they see in the hotel (Annex II, Figure 9). Finally, a service microelement of TV emerged as the tenth top reason for guests' resentment expressed in their reviews (β = 0.228). In this domain, guests routinely complained of non-working, broken or unavailable remote control, small screen size, old TV sets and limited channel choices (Annex II, Figure 10).

5. Discussion

By employing sentiment and semantic network analyses on 109715 hotel reviews with the application of machine learning techniques for big data from online reviews collected from eleven city destinations, this study has exposed and distinguished the causes of low guest satisfaction. Consequently, this research has determined fifty hotel micro-service elements that may commonly induce hotel service failures (Annex I). Simultaneously, a prior mainstream body of research has tended to illuminate groups of hotel service elements or factors rather than particular service elements. In line with the prior studies, this research highlights specific hotel-guest touch points and determines hotel service attributes requiring immediate managerial focus.

In this vein, our study confirms the prior research findings, which revealed significant micro-elements of possible service failures arising from guests' low satisfaction in the domain of the hotel room factor. These micro-elements comprise air conditioning, bed, noise, towels (Luo *et al.*, 2021), wifi (internet) signal quality (Hu *et al.*, 2021), cleanness (Zarezadeh *et al.*, 2022), and hairdryer (Park *et al.*, 2019). However, our research goes beyond the extant corpus of studies as it has revealed a broader range of service micro-elements relevant to service failures under the hotel room factor. Complimentary to prior literature, failing service micro-elements determined by this study comprise room carpet, kettle, table, TV, walls, mattress, room space, shower, tub, sink, smell, fridge, furniture and balcony. Interestingly, the findings of our study point to the hotel room as the key factor in guest satisfaction reduction and, thus, the most frequent area of hotel service failures, as 28 out of 50 service micro-elements are pertinent to the hotel room. In addition, applying the semantic network analysis has made it possible to delve into the context of the hotel service micro-elements. It helped to reveal deeper underlying reasons and roots of guest frustration with particular hotel service micro-elements.

Similarly to Latinopoulos (2020), the completed study confirmed the significance of the hotel location and exterior as a factor of guest satisfaction. On top of that, our study has revealed additional micro-elements relevant to this factor, namely, area, parking, and (hotel) buildings. In the realm of hotel facilities factor, whereas extant research has highlighted solely the impact of their overall functionality on perceived service failures (Ying *et al.*, 2020), this study concretely points to the hotel gym as a micro-element of service failure.

In earlier studies, the role of hotel personnel in contributing to hotel service failures has been widely acknowledged (Nie *et al.*, 2020). Building upon these research findings, this study delves deeper into identifying specific activities that highlight areas where guests often experience frustrations related to hotel personnel. These areas include luggage handling, reception service, shuttle bus service, and animation in resort hotels. Furthermore, our study aligns with the findings of the studies, which highlighted the significance of food and beverage (F&B) in influencing hotel guest satisfaction (Nie *et al.*, 2020; Philips *et al.*, 2017). Surprisingly, our research reveals a relatively lower significance of hotel service micro-elements related to F&B in this respect. According to our findings, only four relevant items, namely 'food', 'drinks', 'breakfast', and 'restaurant', were identified as contributing to service failures. Notably, these items were kept from the top ten micro-elements list, and their positions, based on the log regression coefficient values, were 16, 26, 34, and 39, respectively, in the whole 50 micro-elements range.

5.1. *Theoretical implications*

The completed study contributes to the hotel marketing theory in several ways. First, grounded on the customer's insights from the hotel reviews, this research has zoomed in on malfunctioning hotel services and determined singular service micro-elements representing service failures that cause negative customer experiences and low guest satisfaction. Our research has extended and systemised a set of service micro-elements previously scattered among various studies. We posit that highlighting hotel service factors, or 'attributes', in guest experience research, rather than micro-elements, can lead to a biased understanding of the real reasons for hotel guest dissatisfaction. As previously noted, each hotel service factor typically amalgamates several service micro-elements. Each of them is more or less significant for the hotel guest experience. Hence, we assert that zooming in on the micro-elements is way more significant for the hospitality marketing research than doing so on the hotel service macro factors.

Second, building upon the above point, this study provides a deeper understanding and expands the range of hotel service micro-elements identified in previous literature. In alignment with prior research, our findings corroborate that the internet or Wi-Fi connection is among the top micro-elements that generate negative guest experiences and subsequent dissatisfaction. Additionally, our study highlights the significant role of air conditioning in shaping guests' perceptions of hotel service quality. Notably, Wi-Fi connection and air conditioning emerge as consistent top ten micro-elements in the present research and previous studies.

Moreover, although our findings align with prior research regarding the less significant service micro-elements ranked 11-50 (Table A.1.1.), this study identifies additional service micro-elements that play a crucial role in shaping guest dissatisfaction with the hotel and significantly contribute to the overall understanding of hotel service failures. Our study substantially contributes to the existing literature by uncovering these previously undocumented micro-elements. In summary, this research improves the understanding of the mix of hotel service micro-elements and expands its scope by capturing overlooked factors and highlighting their impact on guest satisfaction. Including these novel findings adds valuable insights to the literature and deepens our comprehension of the multifaceted nature of hotel service failures.

Thirdly, this study extends the understanding of guest dissatisfaction with hotel service quality by shifting the focus towards exploring service micro-elements in eleven diverse city destinations. By examining these micro-elements in randomly selected city destinations, a more comprehensive and holistic understanding of hotel service failures as perceived by hotel guests is achieved. This approach ensures that the findings of this study address the generalisability issue that has been a standard limitation in prior research, which often focused on scrutinising hotel service factors in a single destination. Therefore, this study expands the knowledge regarding the origins of guest dissatisfaction and provides a broader perspective on hotel service failures across different city destinations.

Next, our findings pinpoint specific service micro-elements within the hotel room that are the primary sources of guest complaints. These micro-elements have emerged as significant factors, securing six positions in the list of top-ten service failures (Table A.1.1). Prior research has overlooked the importance of some of these micro-elements relevant to guest satisfaction. Furthermore, our findings highlight that most top service failures are associated with the hotel room, underscoring its paramount role in shaping the guest experience and satisfaction. These insights shed light on the critical influence of the hotel room micro-elements on guest perceptions and emphasise the need for attention and improvement in these areas to enhance overall guest satisfaction.

Finally, in conjunction with the above point, this study has determined and ranked a wide range of hotel service elements and revealed a specific contextual environment pertaining to the failing hotel service micro-elements. Such an approach has yet to be profoundly implemented in the extant body of research. However, it delivers more efficacy in understanding the natural causes of guests' complaints and dissatisfaction with the hotel, as we demonstrated earlier in presenting the results of this study in the previous manuscript sections. Also, evaluating the failing service micro-elements through the semantic network analysis of their context engenders a 360º evaluation of the hotel service micro-elements. It instantly facilitates uncovering the details and contexts behind hotel service failures. The application of this comprehensive approach not only fills a gap in the current body of literature but also enables a more nuanced analysis of hotel service failures. Through this approach, we gain valuable insights into the factors contributing to guest dissatisfaction and provide a more accurate assessment of the prevalent and critical hotel service failures.

5.2. Methodological implications

The present study has several methodological implications. First, grounded in the sentiment analysis and logistic regression procedure synthesis, the present research suggests a multivariate model capable of gauging hotel service components that prompt negative customer emotions and lead to low satisfaction. Hotel guests utilise numeric (rating scores) and textual (review texts) expressions to share their accumulated customer experience. However, quantitative low hotel rating scores cannot alone elucidate the reasons for a negative guest experience. Simultaneously, the hotel reviews dataset, typically containing thousands of lines, belongs to a big data domain and is challenging to analyse nomologically. Hence, developing and availing a workable model solution for hotel review research and analytics is significant.

Secondly, the present research suggests a robust novel methodology for obtaining precise guest insights on the hotel and destination service elements that cause negative customer experiences. The methodology developed for this study is grounded in the instruments suggested and applied in prior studies relevant to the explored topic (Berezina *et al.*, 2016; Kim & Im, 2018; Yadav & Roychoudhury, 2019). As depicted in Figure 1, it encompasses five stages of online guest review data collection, data preprocessing and engineering, sentiment analysis, logistic regression, and semantic network analysis. This developed methodology further develops and advances the accumulated approaches suggested by previous research by (a) introducing the VADER NLP sentiment analysis framework (Hutto & Gilbert, 2014) in hospitality settings, (b) applying a reliable logistic regression procedure to extract singular hotel service elements generating guest frustration, and (c) by implementing semantic network analysis (Israeli *et al.*, 2019; Oh & Kim, 2020) to procure more precise guest insights grounded in the context of poorly operated hotel services. To our knowledge, such blending of the guest review data collection and analysis techniques represents a unique methodological approach to investigating eWOM and hotel guest satisfaction in hospitality research.

5.3. Managerial implications

The accomplished study has several significant implications for hotel management in improving service delivery policies and blueprints. First, in line with the recent research on similar topics relevant to eWOM in hospitality, this study stresses the importance of examining online guest reviews for hotel managers. Online reviews imply a dependable data source that has succeeded conventional data collection techniques implemented in the realm of service evaluation. These canonical models require customer questionnaire-based surveying with several known limitations, including necessary costs, time, and adequate sample size.

Second, our study demonstrates that hotel services can jointly generate guest experience and predispose the tonality of online reviews. That said, poorly operated hotel services commonly engender negative online reviews for the destination. Thereby, it is pivotal for the hotel and DMO managers to work more closely on identifying and improving the poorly operated services negatively evaluated by guests travelling to a tourist destination. Such a combined effort may lead to a rise in the business performance of the destination and its hotels as an entire ecosystem (Henche *et al.*, 2020) through guest satisfaction and loyalty.

Third, our study develops and suggests a practical approach that is effective for examining online reviews and obtaining guest insights to maintain the quality of hotel services. In line with the above-noted implication from the proposed research methodology that is significant for academia, the same approach is fully functional and feasible for implementation by hospitality practitioners. IT developers can produce and market software applications grounded in the algorithm suggested in this study to automate the analytical procedures of eWOM and guest feedback for received customer experiences.

Fourth, the most prominent managerial implication is relevant to hotel services micro-elements representing a particular area of the managers' concern. According to the findings from our research, several hotel service micro-elements may require immediate consideration and improvement. We recommend that hotel managers focus primarily on the hotel service micro-elements that generate higher negative sentiment but simultaneously are manageable. This study determined that such hotel service micro-elements as air conditioning, carpets, gym, kettle, table, luggage, payment, internet, pictures, and TV may reduce and even eliminate guest satisfaction and thus should receive immediate attention from hotel managers. The results of the semantic network analysis employed in this study may help understand the origins of guest dissatisfaction with particular service dysfunctions.

In this domain, hotel managers can undertake several practical steps focused on the critical micro-elements identified by this study. First, regular inspection and maintenance of air conditioning units is paramount. Providing clear instructions on using the thermostat can make a big difference, and upgrading to smart thermostats might be worth considering for the added mutual convenience of the guest and the hotel. Next, keeping up a regular deep-cleaning schedule will keep carpets looking fresh and clean. Quickly addressing any stains or damage and using high-quality, stain-resistant materials can help maintain a spotless appearance. In addition, the hotel must have a working policy regulating carpet replacement when needed.

According to our findings, the hotel gym was another significant managerial concern. In this respect, ensuring the equipment is functional, modern, and well-maintained is essential. The availability of sanitising materials (sprays, wipes, etc.) will encourage guests to clean the equipment before and after use. It also demonstrates that hotel management cares about their guests' health and safety. In upscale hotels, offering complimentary free or paid short fitness classes or personal training sessions can further improve customer experience. Then, hotel housekeeping employees have to check the condition of kettles in guest rooms. Kettles should be cleaned or replaced to ensure they are always in good working condition, given the fragility of this appliance. Various complimentary tea and instant coffee options and clear instructions can please guests more with this service micro-element.

Continuing, hotel managers may consider installing multi-functional tables in the guest rooms. This modern piece of furniture adds flexibility and comfort by allowing guests to use it for dining, working, or relaxing. Regular inspections to fix any wobbling or damage will ensure a seamless experience. The payment process should be straightforward and convenient. Multiple payment options, such as credit/debit cards, mobile payments, and online payments, will cater to different preferences. Clear upfront communication of all charges amid check-in and providing detailed receipts can prevent misunderstandings and build trust with the guests.

Another area for hotel managers to focus on is the luggage storage service. It is essential to ensure ample luggage space in guest rooms and provide assistance with carrying luggage to and from the guest room, as this can significantly improve the guest experience. Upgrading the hotel's internet hardware to provide fast and reliable Wi-Fi signal quality throughout the property is crucial in meeting modern guests' expectations. Offering complimentary high-speed internet access and ensuring the login process is easy and hassle-free is a must in today's digital age.

Next, managers have to ensure that the digital images accompanying the hotel page in OTAs and social media platforms and placed on the hotel web page convey a realistic visualisation of the hotel property without embellishing it. Guests tend to get frustrated when they are misguided by online visualisation as the property needs to meet their expectations. Then, managers should also ensure that guest rooms are equipped with modern, high-definition TV sets that offer a variety of channels in many world spoken languages. Regarding TV, ensuring proper remote control panel work, easy-to-follow TV usage guides, and access to additional content typically make guests' stay more enjoyable. By focusing on these problematic service micro-elements, hotel managers can steer common frustrations and significantly improve the overall guest experience, leading to higher guest satisfaction. Moreover, this study's findings reveal a list of 50 micro-elements noted by guests in their online reviews. Hotel managers should also not overlook the remaining service micro-elements to ensure hotel service quality.

Finally, based on our findings, hotel managers must prioritise the housekeeping and property management departments in their properties. As the majority of identified micro-elements associated with hotel service failures pertain to the hotel room, it is imperative for managers to direct their closest attention to these departments. By focusing on the upkeep and maintenance of the rooms, managers can address the primary sources of service failures and guest frustrations and ensure a more satisfactory experience for their guests.

6. LIMITATIONS, FUTURE RESEARCH, CONCLUSIONS

The present study has some limitations requiring future researchers' consideration when planning their respective studies in the same or similar scope of research. First, the employed SRS procedure facilitated a research scope on hotels operated in eleven tourist destinations. Simultaneously, the literature argues that every tourist destination has specific features and properties (Ostovskaya & Pavlenko, 2018). Therefore, future research may focus on comparative studies between hotels in various city destinations, as scoped by this study, to document differences and similarities in hotel service micro-elements that engender guest dissatisfaction. Moreover, in this regard, the present research is still not without generalisation issues since it evinces poorly operated hotel elements from the perspective of destinations examined in this study. To mitigate the generalisation issues, we recommend that future researchers investigate hotels located in other tourist destinations and determine services affecting guest satisfaction using the methodology suggested by the present study.

Second, in line with the previous research limitation, we suggest that academia contemplate a cultural dimension in their research. Also, we suggest future researchers secure a more even sample in data mining of hotel reviews for cross-cultural studies because the web crawler robot returned a different number of guest reviews per destination in our study due to numerous blank reviews in some destinations and the eventual errors that occurred during the process of reviews collection. As the completed study is relevant to the exploratory type and employs a large sample of qualitative data, such data collection output does not affect the attained findings (Cooksey, 2020).

However, a more even sample will prevent the so-called 'clumping' effect in the sampling procedure, where one descriptive statistic outnumbers other descriptive statistics in the same sample (Peterson, 1975). A strong 'clumping' effect may lead to bias in data distribution and incorrect inferences following the data analysis in comparative studies of city destinations. Thus, we propose that fellow researchers consider a cultural dimension and ensure a balanced sample in their studies. It would generate empirical grounds for comparative and taxonomy studies, delivering intriguing findings.

Lastly, from the standpoint of the research methodology in this study, the literature points to the limitations of the lexicon-based sentiment analysis algorithms employed in our research (Huang *et al.*, 2020). Even though the lexicon-based method is a better solution to gauge sentiment in the guest reviews written in English, as we noted earlier, using the deep machine learning approach employing the CNN technique could benefit future research (Yang *et al.*, 2020). In addition, researchers may investigate opportunities to use alternative machine learning methods for sentiment analysis and causal modelling, including SVM, Naïve Bayes, Random Forest, Stochastic Gradient Boosting, a combination of exploratory factor analysis and linear regression, and others in their studies.

Summarising, this study has revealed and discerned the roots of the negative hotel guest experience by applying sentiment and semantic network analyses with the help of machine learning methods for big data online reviews collected from eleven city destinations. As a result of the application of the methodological procedures, this research extracted fifty hotel service micro-elements that exert guest dissatisfaction with the hotel. The top ten failing micro-elements received a deep examination of their context to spot the origins of guest disappointment. Moreover, the present research suggests a reliable methodology to identify the reasons for low guest satisfaction with hotel services. We hope the developed approach will persuade academia and hotel managers to utilise the suggested methodology in their respective studies to reveal guests' insights on hotel service quality using big data. The findings from this future research will allow managers to detect and revamp poorly operated hotel services in the hotel management domain. Capitalising on the inferences achieved in these future studies will further improve hotel guest satisfaction and, ultimately, the business performance of hospitality organisations.

6. ACKNOWLEDGEMENTS

This research has been conducted within the fundamental research project 2023.001EMPL.VPBUTKOVSKAYA "Application of omnichannel marketing to build successful interaction with consumers in the digital economy" as a part of the HSE Graduate School of Business Research Program in 2023–2025.

7. REFERENCES

Ahn, J., Back, K. J., Barišić, P., & Lee, C. K. (2020). Co-creation and integrated resort experience Croatia: The application of service-dominant logic. *Journal of Destination Marketing and Management*, 17, 100443. https://doi.org/10.1016/j.jdmm.2020.100443

Alam, S., & Yao, N. (2019). The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis. *Computational and. Mathematical Organization Theory*, 25(3), 319-335. https://doi.org/10.1007/s10588-018-9266-8

Alnawas, I., & Hemsley-Brown, J. (2018). The differential effect of cognitive and emotional elements of experience quality on the customer-service provider's relationship. *International Journal of Retail and Distribution Management*, 46(2), 125-147. https://doi.org/10.1108/IJRDM-03-2017-0058

Alrawadieh, Z., & Law, R. (2019). Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers. *International Journal of Culture, Tourism and Hospitality Research*, 13(1), 84-97. https://doi.org/10.1108/IJCTHR-08-2018-0104

Bachleda, C., & Berrada-Fathi, B. (2016). Is negative eWOM more influential than negative pWOM? *Journal of Service. Theory and Practice*, 26(1), 109-132. https://doi.org/10.1108/JSTP-11-2014-0254

Baltes, S., & Ralph, P. (2022). Sampling in software engineering research: A critical review and guidelines. *Empirical Software Engineering*, 27(4), 94. https://doi.org/10.1007/s10664-021-10072-8

Berezina, K., Bilgihan, A., Cobanoglu, C., & Okumus, F. (2016). Understanding satisfied and dissatisfied hotel customers: Text mining of online hotel reviews. *Journal of Hospitality Marketing and Management*, 25(1), 1-24. https://doi.org/10.1080/19368623.2015.983631

Bueno, E. V., Weber, T. B. B., Bomfim, E. L., & Kato, H. T. (2019). Measuring customer experience in service: A systematic review. *The Service Industries Journal*, 39(11-12), 779-798. https://doi.org/10.1080/02642069.2018.1561873

Butkouskaya, V., Romagosa, F., & Noguera, M. (2020). Obstacles to sustainable entrepreneurship amongst tourism students: A gender comparison. *Sustainability*, 12(5), 1812. https://doi.org/10.3390/su12051812

Caceres, R. C., & Paparoidamis, N. G. (2007). Service quality, relationship satisfaction, trust, commitment and business-to-business loyalty. *European Journal of Marketing*, 41(7/8). https://doi.org/10.1108/03090560710752429

Casalo, L. V., Flavian, C., Guinaliu, M., & Ekinci, Y. (2015). Do online hotel rating schemes influence booking behaviors? *International Journal of Hospitality Management*, 49, 28-36. https://doi.org/10.1016/j.ijhm.2015.05.005

Chang, Y. C., Ku, C. H., & Chen, C. H. (2020). Using deep learning and visual analytics to explore hotel reviews and responses. *Tourism Management*, 80, 104129. https://doi.org/10.1016/j.tourman.2020.104129

Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management*, 85, 102356. https://doi.org/10.1016/j.ijhm.2019.102356

Cooksey, R. W. (2020). Illustrating statistical procedures: Finding meaning in quantitative data. *Springer Nature*. https://doi.org/10.1007/978-981-15-2537-7_5

Cuesta-Valiño, P., Bolifa, F., & Núñez-Barriopedro, E. (2020). Sustainable, smart and Muslim-friendly tourist destinations. *Sustainability*, 12(5), 1778. https://doi.org/10.3390/su12051778

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

Henche, B. G., Salvaj, E., & Cuesta-Valiño, P. (2020). A sustainable management model for cultural creative tourism ecosystems. *Sustainability*, *12*(22), 9554. https://doi.org/10.3390/su12229554

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. *Tourism Management*, 61, 43-54. https://doi.org/10.1016/j.tourman.2016.12.022

Hu, F., Li, H., Liu, Y., & Teichert, T. (2020). Optimizing service offerings using asymmetric impact-sentiment-performance analysis. *International Journal of Hospitality Management*, 89, 102557. https://doi.org/10.1016/j.ijhm.2020.102557

Hu, M., & B, L. (2004). Mining opinion features in customer reviews. In *Proceedings of AAAI Conference on Artificial Intelligence*, 4, 755-760.

Hu, X., & Yang, Y. (2021). What makes online reviews helpful in tourism and hospitality? A bare-bones meta-analysis. *Journal of. Hospitality Marketing and Management*, 30(2), 139-158. https://doi.org/10.1080/19368623.2020.1780178

Huang, M., Xie, H., Rao, Y., Feng, J., & Wang, F. L. (2020). Sentiment strength detection with a context-dependent lexicon-based convolutional neural network. *Information Sciences*, 520, 389-399. https://doi.org/10.1016/j.ins.2020.02.026

Huang, S., Liang, L. J., & Choi, H. C. (2022). How we failed in context: A text-mining approach to understanding hotel service failures. *Sustainability*, 14(5), 2675.

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216-225. https://doi.org/10.1609/icwsm.v8i1.14550

Israeli, A. A., Lee, S. A., & Bolden III, E. C. (2019). The impact of escalating service failures and internet addiction behavior on young and older customers' negative eWOM. *Journal of Hospitality and Tourism Management*, 39, 150-157. https://doi.org/10.1016/j.jhtm.2019.04.006

Jang, S., & Lee, J. (2015). The Impact of Hotel Attributes on Online Hotel Reviews: An Analysis of Reviewer Comments. *Journal of Hospitality Marketing & Management*, 24(3), 289-312.

Khan, F. M., Khan, S. A., Shamim, K., Gupta, Y., & Sherwani, S. I. (2022). Analysing customers' reviews and ratings for online food deliveries: A text mining approach. *International Journal of Consumer Studies*, 1, 24. https://doi.org/10.1111/ijcs.12877

Kim, J., & Im, J. (2018). Proposing a missing data method for hospitality research on online customer reviews: An application of imputation approach. *International Journal of Contemporary Hospitality Management*, 30(11), 3250-3267. https://doi.org/10.1108/IJCHM-10-2017-0708

Kirilenko, A. P., Stepchenkova, S. O., Kim, H., & Li, X. (2018). Automated sentiment analysis in tourism: Comparison of approaches. *Journal of Travel Research*, 57(8), 1012-1025. https://doi.org/10.1177/0047287517729757

Kuruuzum, A., & Koksal, C. D. (2010). The impact of service quality on behavioral intention in hospitality industry. *International journal of. Business and Management Studies*, 2(1), 9-15.

Latinopoulos, D. (2020). Analysing the role of urban hotel location in guests' satisfaction. *Anatolia*, 31(4), 636-650. https://doi.org/10.1080/13032917.2020.1808489

Le, T. D., Robinson, L. J., & Dobele, A. R. (2023). eWOM processing from receiver perspective: Conceptualising the relationships. *International Journal of Consumer Studies*, 47(1), 434-450. https://doi.org/10.1111/ijcs.12864

Lee, B. Y., & Park, S. Y. (2019). The role of customer delight and customer equity for loyalty in upscale hotels. *Journal of Hospitality and Tourism Management*, 39, 175-184. https://doi.org/10.1016/j.jhtm.2019.04.003

Lee, C. H., & Cranage, D. A. (2014). Toward understanding consumer processing of negative online word-of-mouth communication: The roles of opinion consensus and organizational response strategies. *Journal of Hospitality and Tourism Research*, 38(3), 330-360 https://doi.org/10.1177/1096348012451455

Lee, M., Cai, Y., DeFranco, A., & Lee, J. (2020). Exploring influential factors affecting guest satisfaction: Big data and business analytics in consumer-generated reviews. *Journal of Hospitality and Tourism Technology*, 11(1), 137-153. https://doi.org/10.1108/JHTT-07-2018-0054

Lee, M., Lee, S., & Koh, Y. (2019). Multisensory experience for enhancing hotel guest experience: Empirical evidence from big data analytics. *International Journal of Contemporary Hospitality Management*, 31(11), 4313-4337. https://doi.org/10.1108/IJCHM-03-2018-0263

Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management*, 71, 337-347. https://doi.org/10.1016/j.tourman.2018.10.004

Luo, J., Huang, S., & Wang, R. (2021). A fine-grained sentiment analysis of online guest reviews of economy hotels in China. *Journal of Hospitality Marketing and Management*, 30(1), 71-95. https://doi.org/10.1080/19368623.2020.1772163

Maimaiti, M., Liu, Y., Luan, H., & Sun, M. (2021). Enriching the transfer learning with pre-trained lexicon embedding for low-resource neural machine translation. *Tsinghua Science and Technology*, *27*(1), 150-163. https://doi.org/10.26599/TST.2020.9010029

Malina, M. A., Nørreklit, H. S., & Selto, F. H. (2011). Lessons learned: Advantages and disadvantages of mixed method research. *Qualitative Research in Accounting and Management*, 8(1), 59-71. https://doi.org/10.1108/11766091111124702

Mariani, M. M., Borghi, M., & Kazakov, S. (2019). The role of language in the online evaluation of hospitality service encounters: An empirical study. *International Journal of Hospitality Management*, 78, 50-58. https://doi.org/10.1016/j.ijhm.2018.11.012

Nie, R. X., Tian, Z. P., Wang, J. Q., & Chin, K. S. (2020). Hotel selection driven by online textual reviews: Applying a semantic partitioned sentiment dictionary and evidence theory. *International Journal of Hospitality Management*, 88, 102495. https://doi.org/10.1016/j.ijhm.2020.102495

Oh, M. M., & Kim, S. S. (2020). Dimensionality of ethnic food fine dining experience: An application of semantic network analysis. *Tourism Management Perspectives*, 35, 100719. https://doi.org/10.1016/j.tmp.2020.100719

Opoku, R. A., Atuobi-Yiadom, N., Chong, C. S., & Abratt, R. (2009). The impact of internal marketing on the perception of service quality in retail banking: A Ghanaian case. *Journal of Financial Services Marketing*, 13(4), 317-329. https://doi.org/10.1057/fsm.2008.26

Ostovskaya, A. A., & Pavlenko, I. G. (2018). Formation of the competitive strategy of development for regional tourist destinations. In F. L. Gaol, N. Filimonova, & V. Maslennikov (Eds.), *Financial and Economic Tools Used in the World Hospitality Industry* (pp. 161-165). CRC Press. https://doi.org/10.1201/9781315148762-31

Padma, P., & Ahn, J. (2020). Guest satisfaction & dissatisfaction in luxury hotels: An application of big data. *International Journal of Hospitality Management*, 84, 102318. https://doi.org/10.1016/j.ijhm.2019.102318

Parasuraman, A., V., Z., & L.L, B. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49(4), 41-50. https://doi.org/10.1177/002224298504900403

Park, H., Kline, S. F., Kim, J., Almanza, B., & Ma, J. (2019). Does hotel cleanliness correlate with surfaces guests contact?". *International Journal of Contemporary Hospitality Management, 31(7), 2933-2950.* https://doi.org/10.1108/IJCHM-02-2018-0105

Park, J., & Lee, B. K. (2021). An opinion-driven decision-support framework for benchmarking hotel service. *Omega*, 103, 102415. https://doi.org/10.1016/j.omega.2021.102415

Peterson, C. H. (1975). The effects of clumping on sample evenness. *The American Naturalist*, 109(967), 373-377. https://doi.org/10.1086/283006

Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: an empirical analysis. *Journal of travel research, 56*(2), 235-249. https://doi.org/10.1177/0047287516636481

Popova, N. (2023, August 3). *Top 100 City Destinations Index 2022 Highlights the Best Performers of the Year*. Euromonitor. https://www.euromonitor.com/article/top-100-city-destinations-index-2022-highlights-the-best-performers-of-the-year

Repovienė, R., & Pažėraitė, A. (2023). Content marketing towards customer value creation. *International Journal of Internet Marketing and Advertising*, *18*(2-3), 263-285. https://doi.org/10.1504/IJIMA.2023.129666

Roy, S. (2018). Effects of customer experience across service types, customer types and time. *Journal of Services Marketing*, 32(4), 400-413. https://doi.org/10.1108/JSM-11-2016-0406

Saleh, F., & Ryan, C. (1992). Conviviality-A source of satisfaction for hotel guests? An application of the Servqual model. In P. Johnson & B.Thomas (Eds.), *Choice and demand in tourism*, 107-122. Mansell Publishing.

Sann, R., Lai, P. C., & Chen, C. T. (2021). Review papers on eWOM: prospects for hospitality industry. *Anatolia*, 32(2), 177-206. https://doi.org/10.1080/13032917.2020.1813183

Shin, S., Du, Q., Ma, Y., Fan, W., & Xiang, Z. (2021). Moderating effects of rating on text and helpfulness in online hotel reviews: An analytical approach. *Journal of Hospitality Marketing and Management*, 30(2), 159-177. https://doi.org/10.1080/19368623.2020.1778596

Son, Y., Kang, K., Choi, I., & Kim, J. (2022). The Determinants of Helpful Hotel Reviews: A Social Influence Perspective. *Sustainability*, *14*(22), 14881. https://doi.org/10.3390/su142214881

Song, Y., Liu, K., Guo, L., Yang, Z., & Jin, M. (2022). Does hotel customer satisfaction change during the COVID-19? A perspective from online reviews. *Journal of Hospitality and Tourism Management*, *51*, 132–138. https://doi.org/10.1016/j.jhtm.2022.02.027

Stamolampros, P., Korfiatis, N., Kourouthanassis, P., & Symitsi, E. (2019). Flying to quality: Cultural influences on online reviews. *Journal of Travel Research*, 58(3), 496-511. https://doi.org/10.1177/0047287518764345

Straka, M., Hajic, J., & Straková, J. (2016). UDPipe: Trainable pipeline for processing CoNLL-U files performing tokenization, morphological analysis, pos tagging and parsing. In *Proceedings of* the *Tenth International Conference on Language Resources and Evaluation, LREC'16*) (pp. 4290-4297).

Straka, M., & Straková, J. (2017). Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipe. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies,* 88-99. https://doi.org/10.18653/v1/K17-3009

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, 68(1), 1-17. https://doi.org/10.1509/jmkg.68.1.1.24036

Vieira, B. M., Borges, A. P., & Vieira, E. P. (2023). The role of social networks for decision-making about tourism destinations. International Journal of Internet Marketing and Advertising, 18(1), 1-27. https://doi.org/10.1504/IJIMA.2023.128148

Wang, D., & Lu, L. (2017). The Impact of Hotel Characteristics on Online Hotel Reviews: Evidence from Chinese Domestic Hotels. *Asia Pacific Journal of Tourism Research*, 22(4), 375-393.

Wang, L., & Kirilenko, A. P. (2021). Do tourists from different countries interpret travel experience with the same feeling? Sentiment analysis of TripAdvisor reviews. In Information and Communication Technologies in Tourism 2021: Proceedings of the ENTER 2021 eTourism Conference, January 19-22, 2021 (pp. 294-301). Springer International Publishing. https://doi.org/10.1007/978-3-030-65785-7_27

Wong, E., Rasoolimanesh, S. M., & Sharif, S. P. (2020). Using online travel agent platforms to determine factors influencing hotel guest satisfaction. *Journal of Hospitality and Tourism Technology*, 11(3), 425-445. https://doi.org/10.1108/JHTT-07-2019-0099

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. https://doi.org/10.1016/j.ijhm.2014.07.007

Yadav, M. L., & Roychoudhury, B. (2019). Effect of trip mode on opinion about hotel aspects: A social media analysis approach. *International Journal of Hospitality Management*, 80, 155-165. https://doi.org/10.1016/j.ijhm.2019.02.002

Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access*, 8, 23522-23530. https://doi.org/10.1109/ACCESS.2020.2969854

Ying, S., Chan, J. H., & Qi, X. (2020). Why are Chinese and North American guests satisfied or dissatisfied with hotels? An application of big data analysis. *International Journal of Contemporary Hospitality Management*, 32(10), 3249-3269. https://doi.org/10.1108/IJCHM-02-2020-0129

Zaibaf, M., Taherikia, F., & Fakharian, M. (2013). Effect of perceived service quality on customer satisfaction in hospitality industry. Gronroos' Service Quality Model Development. *Journal of Hospitality Marketing and Management*, 22(5), 490-504. https://doi.org/10.1080/19368623.2012.670893

Zarezadeh, Z. Z., Rastegar, R., & Xiang, Z. (2022). Big data analytics and hotel guest experience: A critical analysis of the literature. *International Journal of Contemporary Hospitality Management*, 34(6), 2320-2336. https://doi.org/10.1108/IJCHM-10-2021-1293

Zhu, L., Lin, Y., & Cheng, M. (2020). Sentiment and guest satisfaction with peer-to-peer accommodation: When are online ratings more trustworthy? *International Journal of Hospitality Management*, 86, 102369. https://doi.org/10.1016/j.ijhm.2019.102369

APPENDIX A1.1.

Table A.1.1
Hotel service micro-elements defined by logistic regression

	Target variable value = 0 (negative sentizment score)			
	interceipt	**0.068**	**Negative reviews count: *n* = 23135**	**TF-IDF scoring μ:**
	Failing hotel service micro-elements:	**Log. regr. coeff. β value:**		
1	(air) conditioning	0.365	458	0.003
2	carpet	0.358	767	0.003
3	gym	0.356	162	0.001
4	kettle	0.331	373	0.002
5	table	0.307	300	0.001
6	pay	0.300	902	0.004
7	luggage	0.295	352	0.002
8	internet	0.268	306	0.002
9	(online hotel) pictures (images)	0.254	399	0.002
10	TV	0.228	1019	0.003
11	water	0.218	2246	0.007
12	building	0.215	571	0.003
13	elevator	0.213	571	0.003
14	walls	0.213	822	0.003
15	phone	0.212	492	0.002
16	food	0.204	1805	0.007
17	shower	0.201	2747	0.008
18	parking	0.183	515	0.002
19	mattress	0.180	482	0.002
20	stairs	0.179	347	0.001
21	(swimming) pool	0.173	1538	0.008
22	smell	0.170	988	0.005
23	noise	0.160	522	0.003
24	machine	0.159	282	0.002
25	location	0.153	7984	0.019
26	drinks	0.152	186	0.001
27	tub	0.150	205	0.001
28	floors	0.140	315	0.001
29	kitchen	0.140	243	0.001
30	sink	0.137	500	0.002
31	furniture	0.135	565	0.003
32	service	0.129	2541	0.010
33	bed	0.125	3490	0.009
34	breakfast	0.123	4610	0.013
35	area	0,122	1016	0.004
36	doors	0.117	339	0.001
37	room space	0.108	17800	0.022
38	light	0.103	486	0.002
39	restaurant	0.095	918	0.004
40	cleaning	0.083	836	0.004
41	pillows	0.078	436	0.002
42	towels	0.069	1065	0.004
43	bell captain	0.063	353	0.002
44	staff	0.057	6359	0.015
45	animation	0.043	82	0.000
46	reception	0.030	2091	0.006
47	fridge	0.029	367	0.002
48	balcony	0.023	330	0.001
49	slippers	0.020	253	0.002
50	shuttle	0.017	272	0.001

Source: Own elaboration.

APPENDIX A1.2.

Semantic network maps for top ten micro-elements causing hotel service failures

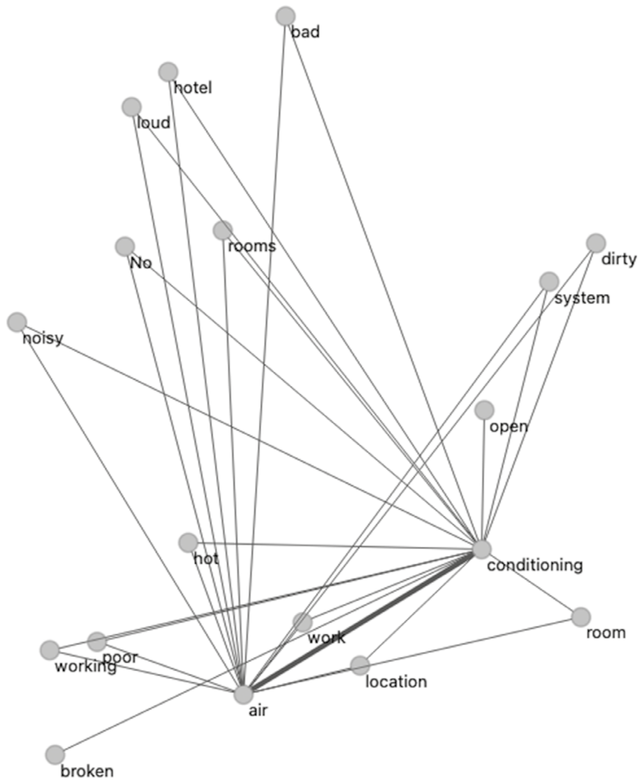

Figure A.1.2.1
Network map: Air conditioning

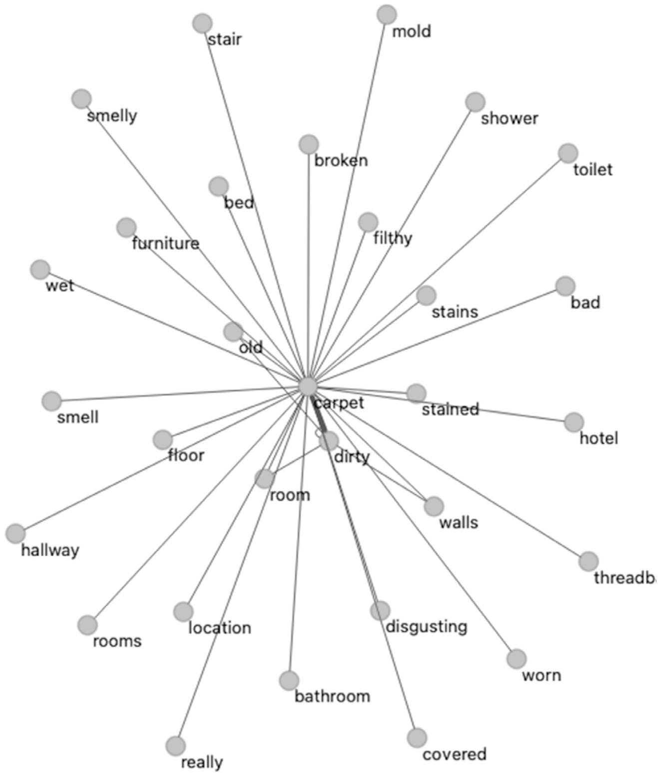

Figure A.1.2.2.
Network map: Carpet

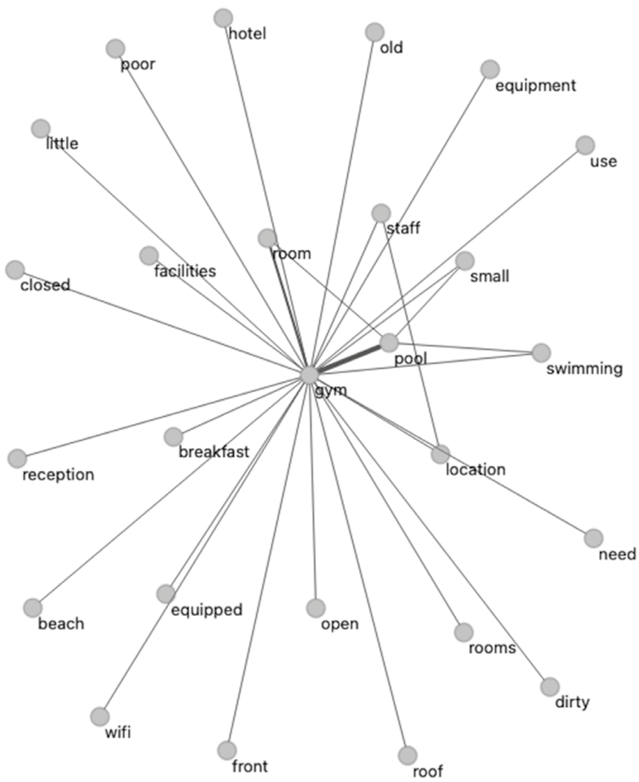

Figure A.1.2.3.
Network map: Gym

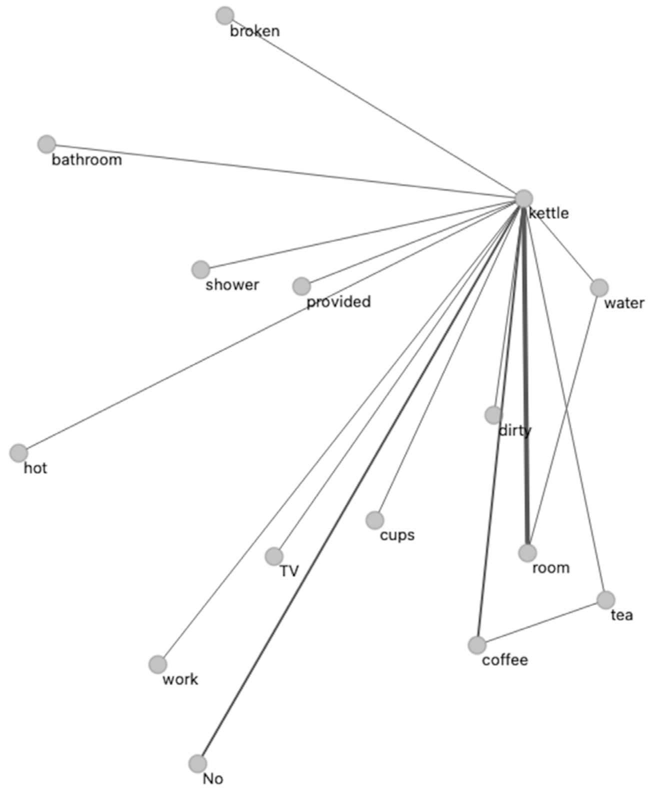

Figure A.1.2.4.
Network map: Kettle

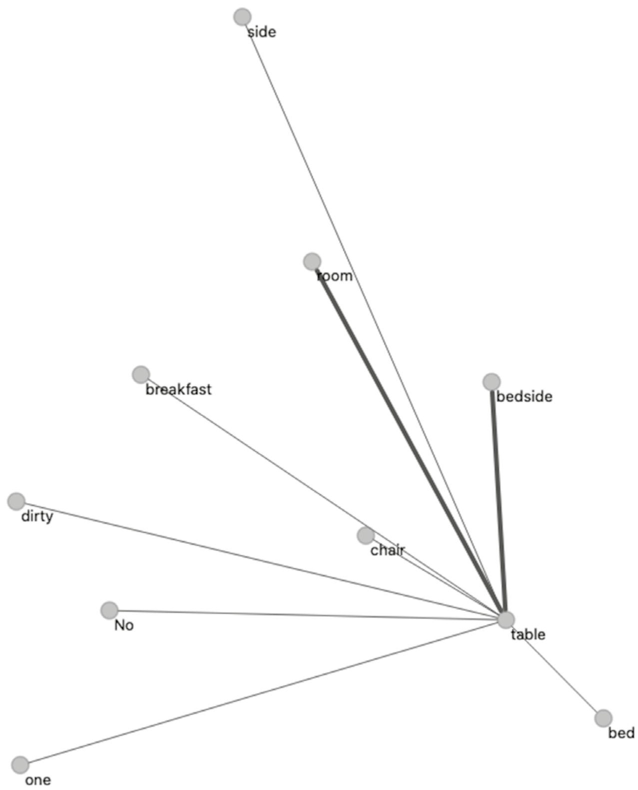

Figure A.1.2.5.
Network map: Table

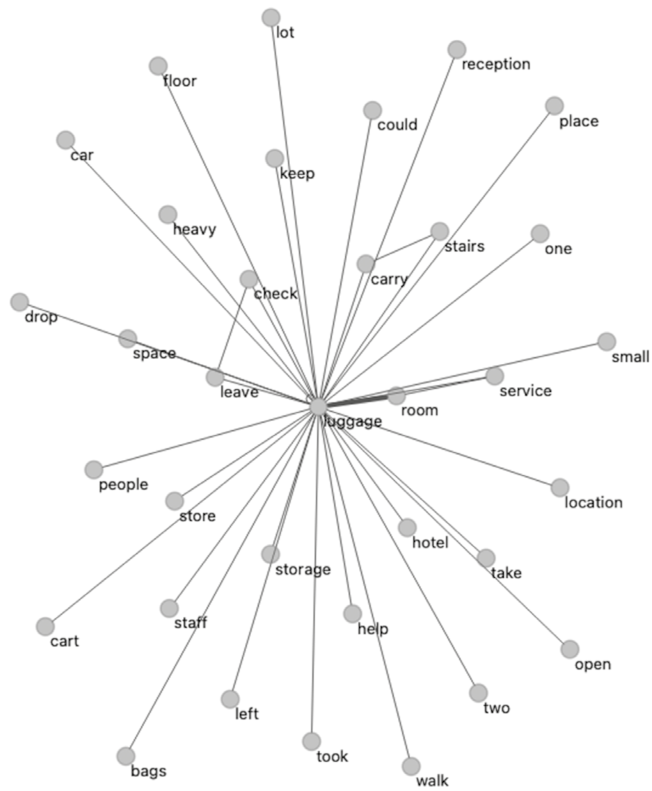

Figure A.1.2.6.
Network map: Luggage

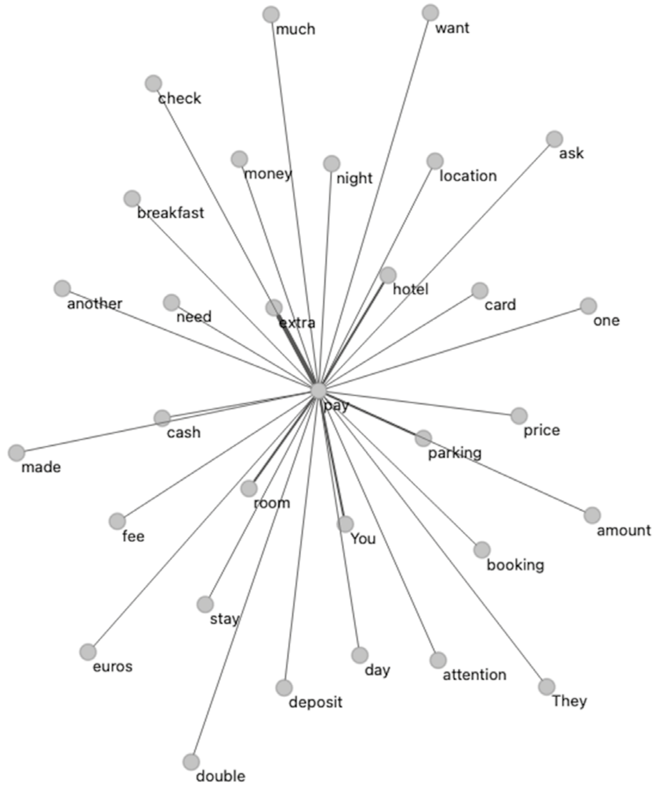

Figure A.1.2.7.
Network map: Pay

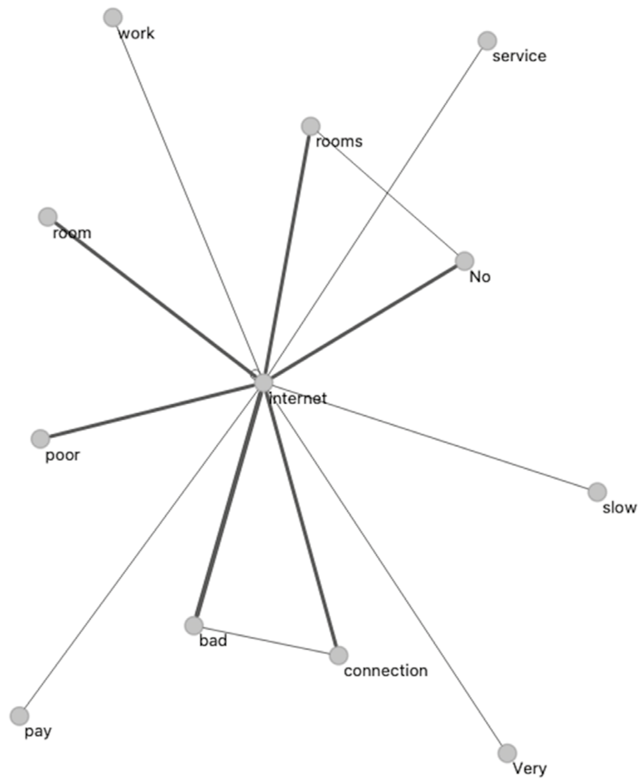

Figure A.1.2.8.
Network map: Internet

Figure A.1.2.9.
Network map: Pictures

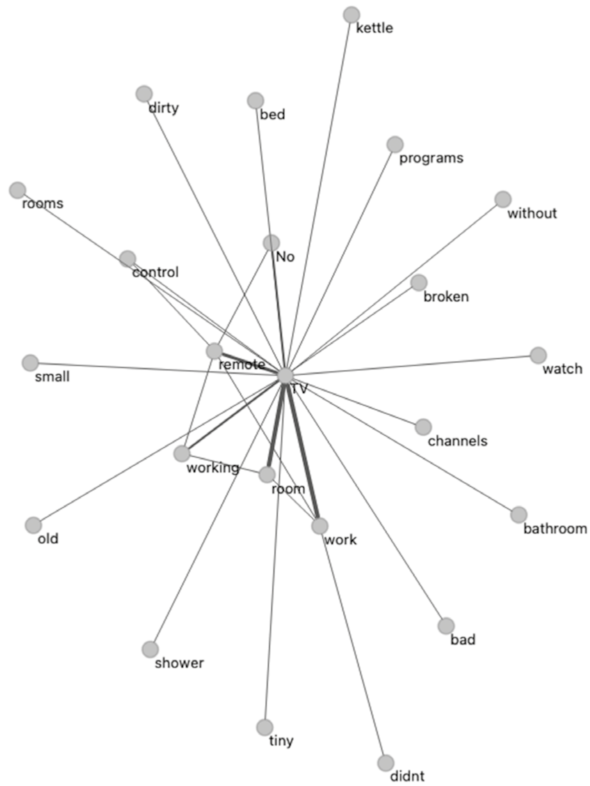

Figure A.1.2.10.
Network map: TV

Source for all figures in Appendix A1.2: Own elaboration.