# The Contribution of Online Reviews for Quality Evaluation of Cultural Tourism Offer: The Experience of Italian Museums


FIRST AUTHOR'S NAME, INITIALS, AND LAST NAME[1]*

First author's affiliation, an Institution with a very long name

SECOND AUTHOR'S NAME, INITIALS, AND LAST NAME

Second author's affiliation, possibly the same institution

THIRD AUTHOR'S NAME, INITIALS, AND LAST NAME

Third author's affiliation, possibly the same institution

FOURTH AUTHOR'S NAME, INITIALS, AND LAST NAME

Fourth author's affiliation, possibly the same institution



In the cultural tourism field, there has been an increasing interest in exploiting data from online reviews adopting data-driven approaches finalized at understanding visitors' perceptions. To date, the sparse studies on the measurement of online perception are mainly based on manual approaches to content analysis and do not compare the visitor and policy maker perspectives. This study addresses this gap, evaluating museum quality dimensions from online reviews of 100 Italian museums over a time-period of one year. Exploiting both a "top-down" approach to the analysis – supervised classification based on policy makers' guidelines – and "bottom-up" approach – unsupervised topic model of online words of reviewers – the resulting quality dimensions are compared, allowing authors to discuss the potential to inform policy making through a user-generated data. Our research contributes to the discussions on the impact that different data analytics approaches have in supporting organizations' decision making processes, by: (1) demonstrating that online reviews can actually provide valuable insights for the evaluation of service quality dimensions defined by decision makers; and (2) showing that a bottom-up approach starting directly from textual expressions in reviews is able to identify further dimensions and quality aspects, that go beyond the typical service-centred analysis performed by businesses and institutions.



**CCS CONCEPTS** •General and reference~Cross-computing tools and techniques ~Empirical studies •General and reference~Document types ~General conference proceedings •Applied computing~Arts and humanities ~Fine arts •Applied computing~Document management and text processing ~Document capture ~Document analysis •Human-centered computing~Collaborative and social computing ~Collaborative and social computing theory, concepts and paradigms ~Social content sharing •Human-centered computing~Collaborative and social computing ~Collaborative and social computing theory, concepts and paradigms ~Social recommendation •Human-centered computing~Collaborative and social computing ~Collaborative and social computing theory, concepts and paradigms ~Social networks •Human-centered computing~Collaborative and social computing ~Collaborative and social computing theory, concepts and paradigms ~Social engineering (social sciences) •Human-centered computing~Collaborative and social computing


[1] * Place the footnote text for the author (if applicable) here.

~Collaborative and social computing theory, concepts and paradigms ~Social media •Human-centered computing~Collaborative and social computing ~Empirical studies in collaborative and social computing

**Additional Keywords and Phrases:** Online user reviews; Visitor perception; Museum quality dimensions; User-driven quality dimensions; Text modelling; Online text analytics

**ACM Reference Format:**

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

## 1 INTRODUCTION

In the cultural tourism field, there has been an increasing interest in exploiting data from online reviews adopting a data-driven approach finalized at understanding visitors' perceptions (e.g., [Sheng and Chen, 2012]; [Schuckert et al., 2015]; [Bi et al., 2019]; [Jia, 2020]). Available studies offer several insights on the expectation of visitors [Sheng and Chen, 2012], opinion of travellers [Hou et al., 2019] or dimensions of service quality [Su and Teng, 2018]. While these analyses are quite diffused for hotels, there is much less evidence on the quality perception for museums. This is mainly due to the absence of clear museum quality dimensions, as opposed to what happens for hotels, where tourist's perception is analysed with reference to predefined dimensions specified by the online review platforms, like cleanliness, location, room, value and service (e.g., [Liu et al., 2017]).

However, museums represent an important area of investigation in the cultural tourism field since they favour the attractiveness of a destination [Giglio et al., 2019] and contribute to the economic development of touristic places [Torre and Scarborough, 2017]. Current literature on the identification of museum quality dimensions from online perceptions of museums visitors through online reviews is limited since the majority of the available contribution investigates museum quality after the onsite experience mainly through customer satisfaction analysis (e.g. [Eleni and Costantine, 2013]; [Moreno Gil et al., 2019]) and surveys (e.g., [Nowacki & Kruczek, 2021]; [Oren et al., 2021]; [Richards et al., 2020]). More recently, manual approaches to the analysis of online reviews have also been used to investigate visitor perception moving beyond customer satisfaction surveys (e.g. [Ferguson et al., 2015], [Craig Wight, 2020]; [Simeon et al., 2017]; [Su & Teng, 2018]). For example, the study by [Su and Teng, 2018] explores the service quality dimension of museums from online reviews, but the analysis is performed manually. Although automatic tools for text analytics have proven to be precious in exploring quality dimensions in various applicative settings (e.g. [Bi et al., 2020]; [Galati and Galati, 2019]), to the best of our knowledge there is still a lack of studies that systematically analyses online reviews for the automatic identification of quality dimensions of museums comparing the expectations of policy makers and the perceptions revealed through the own voice of museum visitors. For instance, [Taecharungroj & Mathayomchan, 2019] analyse of the text of online reviews through a topic modelling technique to identify the most recurrent themes of interest of reviewers, but they target various attractions in the touristic area of Phuket in Thailand instead of focusing on the specific case of museums and they do not compare the results emerging from the online public with the policy maker's perspective. Similar considerations hold for the study of [Kirilenko et al., 2021], that focus on the methodological discussion around the superficial adoption of topic modelling techniques for the automatic detection of online reviewers' thoughts rather than being concerned with the comparison of museum quality dimensions defined by policy makers.

This study addresses this gap by evaluating museum quality dimensions through the analysis of the online reviews of 100 Italian museums over a time-period of one year (2019). The museum quality dimensions evaluated with a "bottom up" approach from the online words of users are compared with the museum quality dimensions evaluated in a "top down" fashion by the policy maker (i.e. the Italian Ministry of Cultural Heritage and Activities and Tourism). The following research questions are here addressed:

- *RQ1:* Which museum quality dimensions can be evaluated in a "top-down" fashion from online reviews?
- *RQ2*: Which museum quality dimensions can be evaluated in a "bottom-up" fashion from online reviews?
- *RQ3*: To what extent do museum quality dimensions evaluated from online reviews in a "bottom-up" fashion differ from those identified in a "top-down" fashion?

The first research question (RQ1) evaluates the museum quality dimensions within the text of online reviews of museums' users through a "top down" approach; indeed, to simulate the traditional approach of policy makers in detecting museum quality dimensions, we develop and apply a keyword-based classifier on top of a list of keywords defined with the policy maker. The second research question (RQ2) explores the museum quality dimensions evaluated by users through a "bottom up" approach for the automatic analysis of the text of the review; specifically, Latent Dirichlet Allocation (LDA) has been adopted to model the user-generated text. The third research question (RQ3) compares the results of the application of the "top down" and "bottom up" approaches to the evaluation of museum quality dimensions from the content of online reviews. This comparison allows the authors to check the (mis)alignment between the "top down" museum quality dimensions expected by the policy makers and the "bottom up" museum quality dimensions directly rising from the own words of online users of museums.

Therefore, from a scientific perspective, the contribution of our work is twofold: (1) to demonstrate that online reviews can actually provide valuable insights for the evaluation of service quality dimensions defined by decision makers; and (2) to show that a bottom-up approach starting directly from textual expressions in reviews is able to identify further dimensions and quality aspects, that go beyond the typical service-centred analysis performed by businesses and institutions.

## 2 DATA COLLECTION AND ANALYTICAL METHODOLOGY

The analysis of the museum quality dimensions from online user-generated reviews is based on the empirical setting of Italian cultural heritage places.

The Italian context is a particularly suited context where to ground cultural studies, for at least two reasons. First, the Italian context is a particularly suited context where to ground cultural studies, since UNESCO recognizes Italy to be one of the countries with the highest density of cultural heritage sites (https://whc.unesco.org/en/list/&order=country#alphaI). Second, in recent years the Italian Ministry for Cultural Heritage and Activities and Tourism has been fostering the digital transformation of tourism as part of its strategic plan of development for 2017-2022 (https://www.turismo.beniculturali.it/en/home-strategic-plan-for-tourism/), pushing cultural institutions to development of digital strategies to promote the cultural heritage and asset, monitor the dynamics of the brand reputation of cultural institutions, and foster the diffusion of digital conversations connected to culture. In line with these directions, since 2018 the authors have been engaged in a project activated by the Italian Ministry for Cultural Heritage and Activities and Tourism with the aim of monitoring the online reputation of a set of 100 Italian public museums, selected by the Ministry itself on the basis of size, geographical distribution and type of collection exhibited. The authors have been engaged in the project for the collection and analysis of museum qualities from the online user-generated contents. More specifically, the Ministry identifieda set of expected qualities of museums and asked the authors to verify the existence of these qualities within the online perception of museums' public.

This allowed the authors to proceed with two parallel approaches. On one side, the expected quality dimensions of museums defined by the policy maker have been searched within online reviewers' texts in a "top-down" fashion, classifying online reviews in a supervised way according to the policy maker's expectations. On the other side, the authors also followed a "bottom-up" approach to analyse in a data-driven and unsupervised fashion the text of online reviews to identify the museum quality dimensions directly from the words of online users. The results of the "top-down" and "bottom-up" analyses of reviews presented within this paper allows the authors to discuss the potentialities of the bottom-up approach in grasping the perceptions of quality dimensions directly from the own words of users.

Below we detail the processes of data collection and enrichment (Section 2.1), describe data (Section 2.2) and the details of the data analytics procedures (Section 2.2), conducted with the top-down approach (Section 2.2.1), the bottom-up approach (Section 2.2.2) and comparing the two approaches (Section 2.2.3).

### *2.1 Data Collection & Enrichment*

We built a customized automated system and data analysis pipeline that collects, cleanse, and processes the online reviews of museums coming from the TripAdvisor webpages of the 100 Italian public museums selected by the Italian Ministry of Cultural Heritage and Activities and Tourism.

For each of these museums, we manually identified the TripAdvisor webpages and verified the credibility of the web sources directly with museum managers. We then implemented the automated and scheduled data collection system, storing the online user-generated reviews in a document-based storage solution. This allowed the incremental update of the collections and enabled a daily monitoring of the online reputation of museums. Thanks to the data collection system, we were able to collect 47,993 online reviews published along 2019 on the TripAdvisor pages of the 100 Italian museums of interest for the study.

Once collected, the online reviews of museums were enriched following this 2-steps procedure:

- *Language detection*: the language of each online review was identified using a pre-trained Google model (implemented in the *langdetect* Python package, see also https://github.com/Mimino666/langdetect) with the help of an external service (Dandelion API, see https://dandelion.eu/docs/api/datatxt/li/) to ensure the consistency of the result. The precision of the state-of-the-art language detection techniques are over 99% for 53 different languages (the original project and performance results are described at https://code.google.com/archive/p/language-detection/ and in the original project presentation at https://www.slideshare.net/shuyo/language-detection-library-for-java).
- *Sentiment Analysis*: focusing on 14,250 online reviews of museums automatically recognized to have been written in Italian, the Italian text of online reviews was then provided as input to calculate reviews' sentiment. To grant maximum quality in the analysis, we compared four methods:
  - a custom javascript implementation of AFINN, one of the simplest techniques for sentiment analysis (https://github.com/fnielsen/afinn), which associates a score between -1 and +1 and a label (negative, neutral, positive) to each review;
  - an external service, named Dandelion (see https://dandelion.eu/docs/api/datatxt/sent/), which associates a score between -1 and +1 and a label (negative, neutral, positive);
  - a combination of the previous two methods, aimed at resolving conflicts between the scores provided by two methods separately
  - a language model method, namely BERT (Devlin et al. 2018), published by researchers at Google AI Language.

  Since the accuracy for the Dandelion classifier on the Italian reviews of museums under investigation was 41% and that of the AFINN-based model was 12%, we combined these two methods to make the final result of sentiment classification more consistent. However, their combination reached an accuracy of just 42%. Therefore, we decided to use BERT, a method that caused a stir in the Machine Learning community achieving state-of-the-art results in a wide variety of NLP tasks. Since we were specifically interested in the detection of sentiment of Italian reviews, we selected a BERT model pre-trained specifically on social media contents (i.e. Twitter) written in Italian language (https://huggingface.co/neuraly/bert-base-italian-cased-sentiment). Thanks to this choice, the model was already prepared for our scenario, allowing us to avoid further training of the model. We evaluated the method over sentiment labels associated with a random sample of 1000 Italian reviews. As a result, we reached an accuracy of 86% with this technique, which we selected as sentiment algorithm used for the analysis.

To further validate the quality of the collected data and the consistency of the data analytics, results were also displayed in a dashboard. Granting a real-time access to the dashboard, policy makers and museum managers were able to visualize, explore and monitor the online reputation of museums on TripAdvisor and on other online channels, like online news websites and social media platforms, such as Facebook, Instagram and Twitter. The real-time access to the dashboard also fostered a frequent communication among policy makers, museum managers, and researchers, thus allowing continuous quality validation of the analyses and results provided.

Overall, the data collection and enrichment procedure results in a dataset of 14,250 online reviews of museums automatically collected from TripAdvisor and for which the language has automatically been recognized to be Italian, enriched with a sentiment score provided by the model recognized as state-of-the art model for NLP tasks.

### *2.2 Data Description*

The analysis of the Italian reviews of museums shows a seasonality in the amount of reviews and in the quantitative evaluation (i.e., rating) and qualitative evaluation (i.e., sentiment) of the quality of museum visits.

Looking at the overall number of reviews (Figure 1), there is a peak in spring and late summer, with 1,854 reviews in April and 1,506 in August. This can be connected to school trips and visits by foreign tourists that prefer spring and late summer to visit Italy [Guizzardi and Mazzocchi, 2010]. On the contrary, the amount of reviews decreases in winter, with particular reference to the month of December (778 reviews) and February (953 reviews).

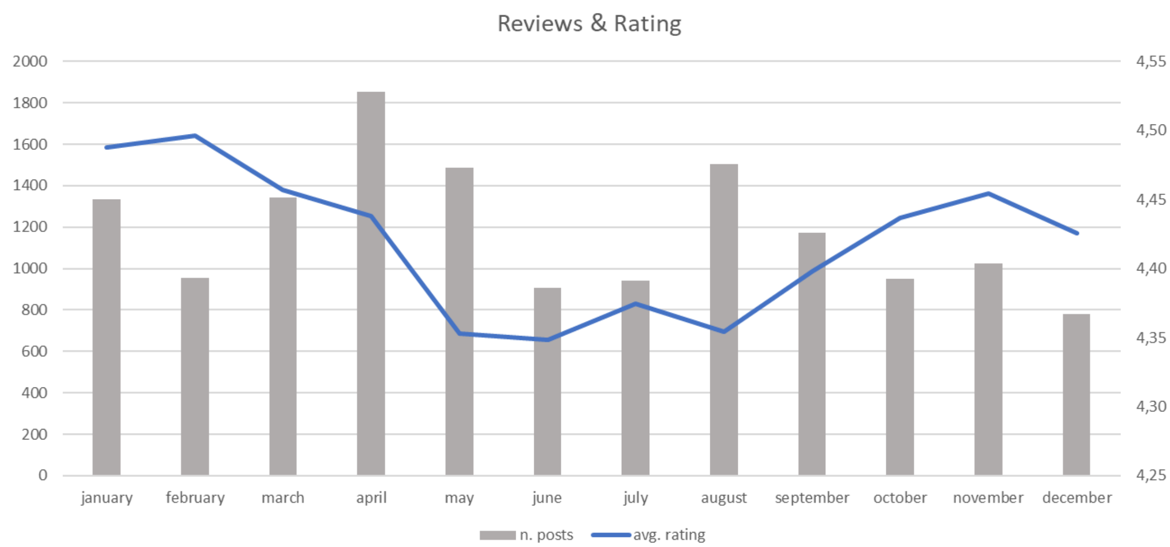


Figure 1 Evolution of number of Italian reviews and average rating in 2019, monthly data.

Focusing the attention specifically on the quantitative evaluation of the quality of visit captured by the rating of online reviews, the starting point was the analysis of the quantitative value of the online rating offered by TripAdvisor. Results show that the average annual rating of museums' reviews is 4.42 out of 5, with limited variability of the monthly average ratings, since values fall between 4.35 and 4.50 stars. However, the distribution of ratings month by month shows values closer to 4.50 at the beginning of the year (January-March) than during late spring and summer, a period in which the average rating achieves the minimum values (4.35 stars over 5 in August). Therefore, we observe a behaviour out of phase between the number of reviews and their ratings: periods like spring and late summer during which the number of reviews presents picks, show valleys in the average ratings while the behaviour is reversed in autumn and winter. This could be connected to a more positive perception of museums when there is less crowd, hence in quiet periods as winter.

The qualitative evaluation captured through the sentiment of online reviewers, confirms the observations previously made on the quantitative evaluations (Figure 2). Indeed, the average annual sentiment score of museums' reviews is 0.7499, on a range from -1 (negative) to +1 (positive). We noted not only an extremely high value of the average sentiment of reviews month by month but also little variability, with a minimum value of 0.7009 in May and a maximum value of 0.8099 in February. As observed for the trend of the monthly rating, though the values of sentiment remain always extremely positive, at the beginning of the year (0.7855 average January-March) museums record higher values than the one obtained in late spring and summer (0.7331 May-September). As observed for the trend of the monthly rating, also the sentiment of reviews is out of phase with the number of reviews, registering higher values when the number of reviews is lower and vice versa.

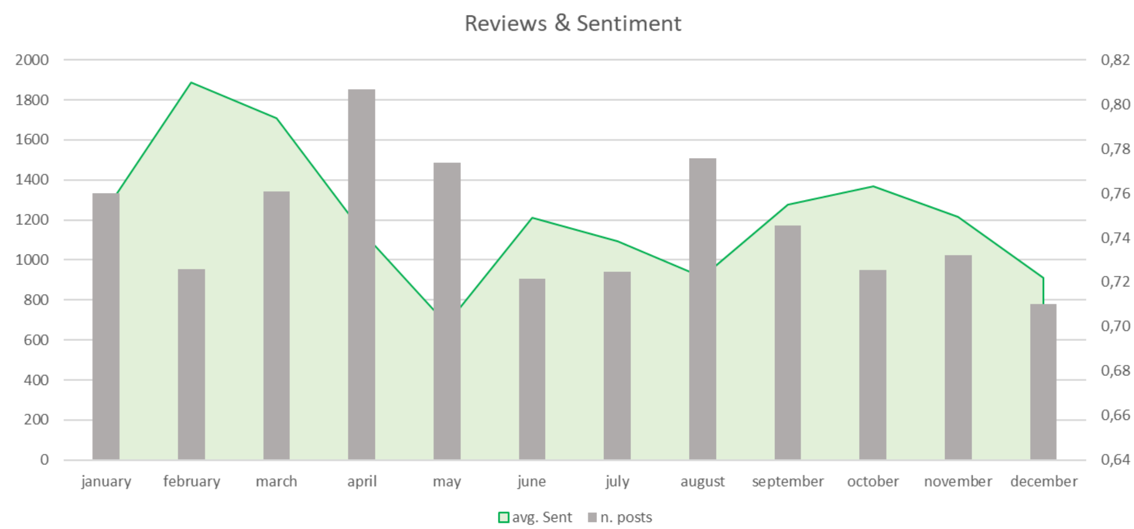


Figure 2 Evolution of number of Italian reviews and average sentiment in 2019, monthly data.

### *2.3 Data Analysis*

To replicate the "top-down" and "bottom-up" approach for the identification of museum quality dimensions, we selected respectively (i) a supervised keyword-based classifier for text of reviews, based on a set of keywords defined by the policy maker (Section 2.3.1) and (ii) an unsupervised statistical topic model for the automatic modelling of latent themes of discussions within the online user-generated text of reviews (Section 2.3.2). Comparing the "top down" and "bottom up" approaches allows us to underline the differences in the supervised and unsupervised models adopted (Section 2.3.3), preparing the reader for the interpretation and discussion of the results, which show a (mis)alignment between the "top-down" museum quality dimensions expected by the policy makers and the "bottom-up" museum quality dimensions directly rising from the own words of online users of museums.

#### *2.3.1 "Top-down" Approach*

This section discusses the results obtained from the adoption of the "top-down" approach (Table 1), here intended with reference to a set of quality dimensions defined by the policy maker.

Table 1 “Top-down” quality dimensions of museums, derived from the directions of the policy maker. Keywords are translated in English to increase readability and comprehension, but the algorithm uses the original words in Italian.

| *Policy maker' standards* | *"Top-down" Quality Dimension* | *Set of keywords* |
|---|---|---|
| *Perception of museum user/visitor towards reception services, i.e., friendliness and professionalism of the staff, hostesses, stewards, security guards, cleaners, ticket office, presence of queues, queues, crowds, waiting, flow management, route management, etc.*<br>*Perception of museum user/visitor towards costs of basic and additional services, i.e., costs incurred / to be incurred* | Ticketing & Welcoming | *Free, sliding, queue, free, ticket, throng, crowd, wait, entrance, steward, checkout, cost* |
| *Perception of museum user/visitor towards museum's physical location*<br>*Perception of museum user/visitor towards accessibility, e.g., lifts, platforms / slides for people with disabilities, transport to reach the place, parking lots* | Space | *Restoration, dirt, external / externally, intern, enter, unkempt, access, imposing, out* |
| *Perception of museum user/visitor towards use of spaces, e.g., halls, exhibitions, set-ups, lighting, signage, aesthetics, exhibition, captions, tour itinerary, cleaning, indoor locations, outdoor locations* | Comfort | *lighting, degraded, cured, held* |
| *Perception of museum user/visitor towards organized activities and events, e.g., exhibitions, shows, presentations, experiences, workshops, educational activities, events, etc.* | Activities | *guided tour, show, event* |
| *Perception of museum user/visitor towards usability, usefulness, completeness and quality of the applications and content of websites and social accounts, i.e., clarity, completeness and exhaustiveness of the information on the institution available online by consulting the institutional websites and social media / network accounts; quality and usability of the applications available online for booking and purchasing tickets for visiting institutes / exhibitions and the like; website compatibility with mobile devices* | Communication | *Guide, videoguide, audioguide, package, audio, video, booking, indication, itineraries* |

In our empirical validation setting, the policy maker is represented by the Italian Ministry of Cultural Heritage and Activities and Tourism, who introduced in 2018 a set of quality standards for public museums with a Ministerial Decree (Ministerial Decree Nr. 113, 21/02/2018). These standards delineate a list of relevant aspects for which museums are held accountable (see *Policy maker' standards* in Table 1).Based on this list and on the direct interaction with policy makers, we were able to group in a “top-down” fashion the policy maker's aspects in five quality dimensions, i.e. Ticketing & Welcoming, Space, Comfort, Activities, and Communication (see *“Top-down” quality dimensions* in Table 1). Beyond verifying the consistency of these dimensions, the policy maker was also asked to verify the association to these five dimensions of a set of keywords, that we automatically extracted among the most frequently used words of online reviewers of museums. The list of keywords remaining after the policy maker's validation (see *Set of keywords* in Table 1) was then used to build a keyword-based classifier for the classification of online reviews' text into each of the five dimensions.

We choose a keyword-based classifier for the classification of online reviews' text to simulate the “top-down” approach adopted by museum managers and policy makers, that is to predefine specific categories of museum quality dimensions to be evaluated and search for occurrences of these categories in the opinion of the general public. The automatic keyword-based classifier of online reviews' text developed and implemented in this paper can then be thought as a proxy of the procedure museum managers and policy makers usually manually perform for evaluating the existence of museum quality dimensions within online reviews. Indeed, given a set of specific museum qualities defined by policy makers, the supervised keyword-based classifier acts as the museum manager, searching for occurrences of these categories within public opinion. In this sense, the implementation of the keyword-based classifier for the text of online reviews can be interpreted as the automated version of the manual check performed by museum managers on the content of online reviews.

Focusing on the 14,250 Italian reviews of museums, we built a non-overlapping multiclass keyword-based classifier to assigns reviews to the five classes, i.e. Ticketing & Welcoming, Space, Comfort, Activities, and Communication, based on presence or absence of specific keywords in the text of the review (Table 1). Since the five "top-down" quality dimensions are not mutually exclusive nor exhaustive, one review can simultaneously be associated to more than one class or to none (Figure 3). In the latter case, we label such a review with the term "Other Aspects" to underline that the review is not connected to any of the quality categories defined with the policy maker.

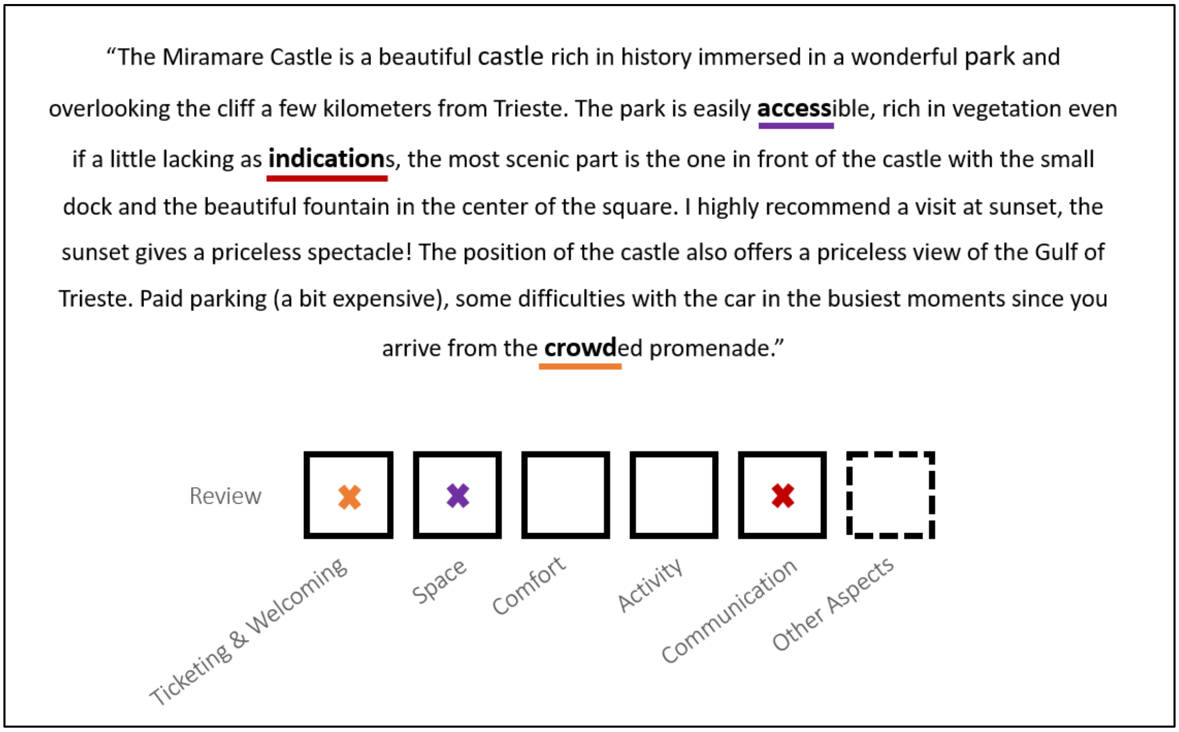


Figure 3 Example of the result of the application of the "top-down" approach for the identification of the museum quality dimensions within online reviews, based on a non-overlapping multiclass keyword-based classifier. The review is classified in three out of the five predefined museum quality dimensions.

To select the classification algorithm, we followed the same methodological steps adopted for the selection of the most suitable sentiment algorithm (Section 2.1). In particular, we compared the keyword-based classifier with the BERT algorithm for Italian language already presented in Section 2.1. To test the algorithmic performances, we randomly sampled 1000 Italian online reviews of museums and manually screened their text to assign value 1 to each "top-down" category whenever the text of the review was indeed addressing the aspects connected to the category. We split with stratified k-fold cross validation the manually labelled data to obtain a training set of 800 reviews and a testing set of 200 reviews. Also thanks to the frequent monitoring of online reviews supported by the dashboard we developed (Section 2.1), we already expected highly unbalanced data prior to the application of the classifier. This expectation was confirmed not only in the randomly sampled reviews used to test the performances but also in the overall dataset considered for the analyses, as shown in the results (see also Section 3.2).

This reflected on the performances of algorithms: the keyword-based method obtained an average accuracy of 80% and recall of 50% among the five classes (Table 2), while the BERT method obtained 88.2% accuracy and 58% of recall. Notwithstanding the slightly higher performances of BERT, we selected the keyword-based classifier over BERT method, since the latter was seriously affected by the unbalanced nature of data being unable to predict three out of the five categories.

Table 2 Performance of keyword-based classifier for each of the “top-down” quality dimensions and average across categories.

| *“Top-down” quality dimension* | *Accuracy* |
|---|---|
| *Ticketing & Welcoming* | *72%* |
| *Space* | *76%* |
| *Comfort* | *80%* |
| *Activity* | *94%* |
| *Communication* | *75%* |
| ***Average*** | ***80%*** |

### *2.3.2 “Bottom-up” Approach*

To automatically exploit in a “bottom-up” fashion the perspective of visitors on museums’ quality dimensions, we resort to the application of an unsupervised model for the text of online reviews of museums based on topic modelling. The statistical detection of latent topics of discussions among reviewers’ words has been performed applying Latent Dirichlet Allocation (LDA) (Blei, 2012) model to the text of reviews for two main reasons.

First, this generative probabilistic model entails the peculiar characteristics of Bayesian models to be highly flexible to the specific domain of application. Second, this method allows us to detect hidden structures within text of online reviews in terms of semantically similar groups of words, namely latent topics of discussion, that we interpret as latent museum quality dimensions hidden within the own words of online reviewers of museums. Indeed, by choosing the LDA procedure to define a set of topic-based quality dimensions of museums, we are simulating the process of visitors in evaluating the quality of museums: without even acknowledging how many or which could be the dimensions of quality of a museum, visitors evaluate their visit emphasizing various aspects, which hinder the perceived quality of museums.

Focusing on the 14,250 Italian reviews, we implemented LDA in R environment (https://cran.r-project.org/). Resorting to *tm* and *SnowballC* packages, the pre-processing phase consisted in converting text to lowercase, removing particular characters (e.g., emojis, URLs, punctuation and numbers), excluding language-specific and context-specific stopwords (i.e., *roma, colosseo, pantheon, pantheum, phantheon, pompei, firenze*) and reducing the grammatical forms of words through Porter’s stemming algorithm. Then, the four metrics proposed by [Griffiths & Steyvers, 2004], [Juan et al., 2009], [Arun et al. ,2010] and [Deveaud et al.,2014] and implemented in *FindTopicsNumber* function, of package *ldatuning* have been used to select an appropriate number of topics between 2 and 30. Resorting to the elbow criterion, each of the plausible configurations of latent topics of discussion identified has been interpreted considering the 30 words with highest probability values in the per-topic word distributions and the reviews with the highest probability values in the per-document topic distributions.

To increase the interpretability of the selected LDA model, we further grouped the resulting latent topics of discussion into three main “bottom-up” dimensions of museum quality, that we interpreted as *Museum Cultural Heritage*, *Personal Experience*, and *Museum Services* (detailed description in Section 3.2). Thanks to the adoption of the LDA probabilistic topic model, the “bottom-up” representation of each review is a mixture of three “bottom-up” quality dimensions of museums, where the probability of observing a specific quality dimension is given by the emphasis given by the reviewer to the specific museum quality dimensions (Figure 4).

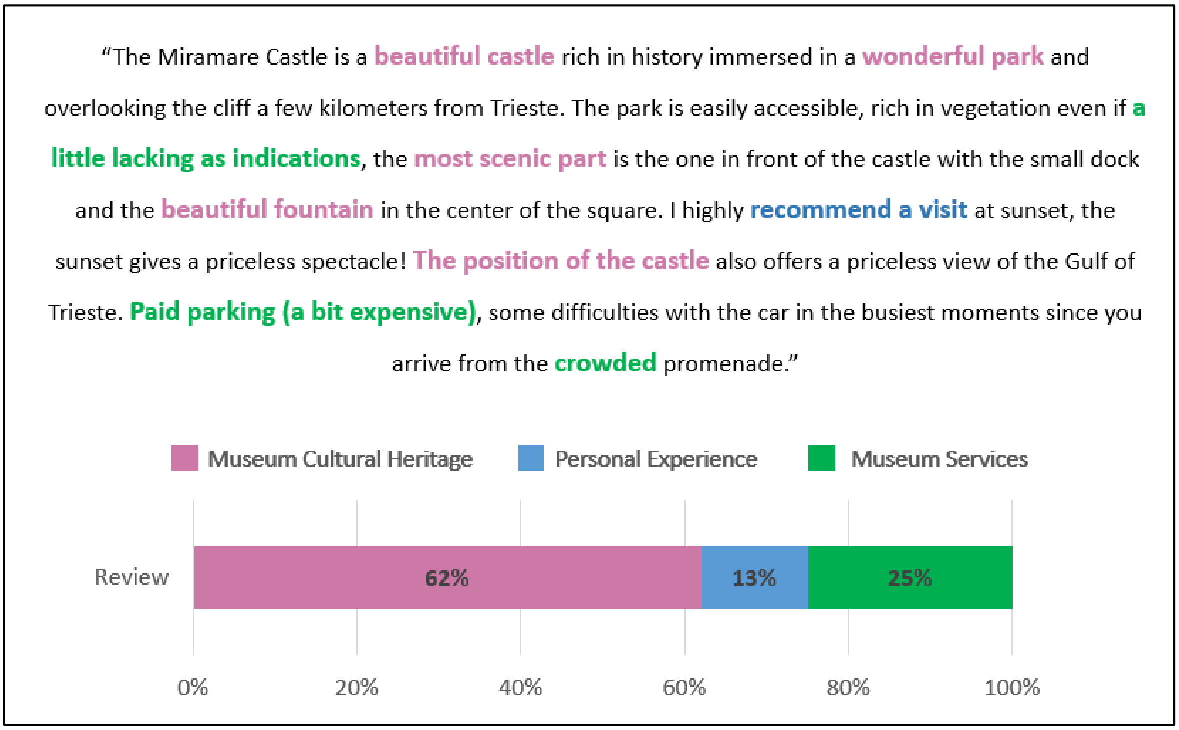


Figure 4 Example of the result of the application of the "bottom-up" approach for the identification of the museum quality dimensions within online reviews, based on a LDA topic model. The review is a mixture of the three "bottom-up" museum quality dimensions, where each proportion depends on the emphasis with which the reviewer discusses the corresponding museum quality.

*2.3.3 Comparison of "Top-down" and "Bottom-up" approaches*

From a modelling perspective, the "top-down" and "bottom-up" approaches differ across at least five features: perspective, categorization, training, interpretation and representation (Table 3).

In terms of perspective and categorization, the two methods are complementary. The "top-down" approach is a supervised model that simulates the behaviour of the policy maker, who defines a set of quality dimensions which should be evaluated by museum managers within the opinion of the general public. The supervised model based on the specific set of keywords defined by the policy maker is hence just an automated version of the actual process already adopted. On the contrary, the "bottom-up" approach is an unsupervised topic-based model that simulates the museum visitor's perspectives, identifying quality dimensions of museums from the latent dimensions of museum visit detected within the own words of online reviewers. Following this logic, this approach results in a set of probable aspects not defined a priori, but that need to be interpreted.

Since these aspects are hidden within the words of visitors, prior to the analysis there is no clear indication on the number of dimensions to be searched for nor on the specific contents to be searched for, requiring hence a certain effort to interpret the "bottom-up" results. On the contrary, once the keywords and the categories have been defined, the results of the "top-down" approach are of immediate interpretation since each review is either associated or not to each of the specific categories according to the identification of the specific words within its text.

From the point of view of the implementation, the two models have different requirements in terms of training. While the "top-down" approach needs to learn how to search for categories with a training set, the "bottom-up" approach learns the hidden structures directly from data without requiring training.

The two models also differ in terms of output and hence of representation of each review. With the “top-down” approach, each review is represented as a sequence of length given by the number of categories of the classifier, where each entry indicates in a binary way whether the specific categorization has been found to be present in the text of reviews or not. With the “bottom-up” approach, each review is still a sequence of length given by the number of dimensions retrieved, but with each entry indicating the probability of referring to each specific dimension. Compared to the “top-down”, the “bottom-up” representation implies that each review has a potentially non-zero value of discussing each of the dimensions. Also, the “bottom-up” representation allows to rank the quality dimensions within each review from the most discussed to the least discussed, and to rank reviews according to their propensity to discuss specific quality dimensions, property not captured by the “top-down” representation.

Table 3 Comparison of top-down and bottom-up approaches to identification of museum service quality dimensions.

| | *“Top-down” model* | *“Bottom-up” model* |
|---|---|---|
| *Perspective* | *Policy Maker / Museum Manager* | *Reviewer / User / Visitor* |
| *Categorization* | *Supervised (keyword-based)* | *Unsupervised (topic-based)* |
| *Training* | *Required* | *Not required* |
| *Results interpretation* | *Not required* | *Required* |
| *Representation* | *Non-overlapping multiclass classifier output* | *Topic-based probability distribution* |

## 3 RESULTS AND DISCUSSION

This section reports the results of the analysis of the 14,250 Italian TripAdvisor reviews received in 2019 by the 100 Italian public museums selected by the Italian Ministry of Cultural Heritage and Activities and Tourism. Results are here presented to answer each of the three research questions and to highlight the value that museum experts can extract from the automatic modelling of text of online reviews. The comparison of the policy maker perspective - grasped through the “top-down” approach (Section 3.1) - and of the online visitors’ perspectives - grasped through the “bottom-up” approach (Section 3.2) - shows a (mis)alignment between the “top down” and “bottom-up” museum quality dimensions (Section 3.3) and enriches the discussion of the methodological differences between the two models (Section 2.3.3).

### 3.1 *RQ 1. Which museum quality dimensions can be evaluated in a “top-down” fashion from online reviews?*

The automatic supervised non-overlapping multiclass keyword-based classifier developed and implemented in this paper is based on the five classes defined by the policy maker (Table 4).

The application of the ″top-down″ approach resulted in a limited amount of reviews classified within the aforementioned five categories (Figure 5), with the 63% of the analysed reviews not assigned to any of the five dimensions identified by the policy maker and therefore labelled by us as belonging to the *Other Aspects* category.

Looking closer at the content of these reviews, we found that the ″top-down″ approach supported the evaluation of specific dimensions of interest for the policy maker, but failed in detecting the many other aspects of interest of reviewers of museums, which go beyond the set of keywords predefined by the policy maker. While the five classes defined by the policy maker are limited to *services* offered by the museum, like ticketing, communication and activities, the museum public does not necessarily underline only these service-related aspects but rather address additional aspects.

To further explore the perspective of online reviewers and the misalignment with respect to the policy maker’s perspective, we investigated in an automated “bottom-up” fashion the museum quality dimensions hidden within the content of online reviews, as described in the next section.

Table 4 Short description of “top-down” classes and examples of excerpts of reviews classified in the corresponding class.

| *“Top-down” class name* | *Short description* | *Excerpt of review classified* |
|---|---|---|
| *Ticketing & Welcoming* | *aspects related to surveillance, welcoming, fees and costs, such as ticketing, queueing and crowding* | *… Paid parking (a bit expensive), some difficulties with the car in the busiest moments since you arrive from the crowded promenade.* |
| *Space* | *aspects related to the physical characteristics of the museum, such as the location of the museum and its accessibility* | *… beautiful castle rich in history immersed in a wonderful park and overlooking the cliff a few kilometers from Trieste. The park is easily accessible, rich in vegetation ...* |
| *Comfort* | *aspects related to the equipment of museums' exhibitions, such as lighting, cleanliness or maintenance* | *... the rooms with well-kept furnishings, paintings, furnishings that are well preserved and repaired from tampering ...* |
| *Activity* | *aspects related to events organized by museums, such as guided tours and temporary exhibitions* | *... The advice is to book a guided tour of at least 4 h, as we did, and you will not regret it, as a shorter time is really small ...* |
| *Communication* | *aspects related to information offered to the public onsite or through online channels, such as physical signposts and audio-guides* | *… even if a little lacking as indications, the most scenic part is the one in front of the castle with the small dock and the beautiful fountain in the center of the square ...* |

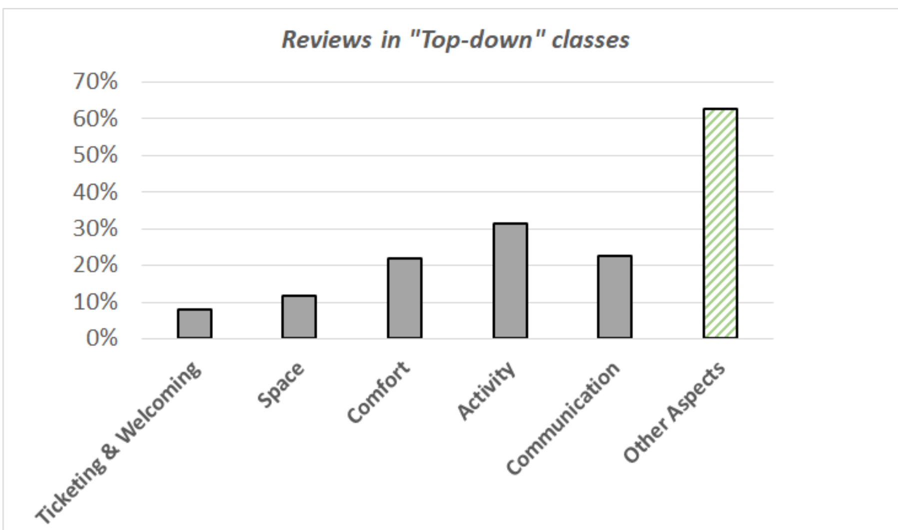


Figure 5 Proportion of Italian reviews that have been classified in each of the “top-down” class identified by policy makers (solid bars) and proportion of Italian reviews not associated to any “top-down” category (striped bar). Notice that percentages do not sum up to 100% due to adoption of a non-overlapping multiclassifier.

### *3.2 RQ 2. What museum quality dimensions can be evaluated in a “bottom-up” fashion from online reviews?*

This section provides the results of the “bottom-up” approach based on the application of an LDA model to the same dataset analysed in the previous section. This analysis required us to interpret the 13 latent topics resulting from the analysis of topics of online reviewers, that we further interpreted as representing three “bottom-up” quality dimensions (Table 5):

- *Museum Cultural Heritage (6 latent topics):* museum reviews address with an average probability of 46% aspects connected to artistic collection of the museum, including comments on exhibitions, findings and artworks, but also considerations on museum's history and tradition and descriptions of museums' buildings, facades, churches and castles.
- *Personal Experience (4 latent topics)*: museum reviews address with an average probability of 31% emotional aspects associated with their personal experiences. This includes comments connected to the "wow effect" of the visit, praises to the majesty and beauty and suggestions to visit at least once in a lifetime the heritage site. Other aspects addressed are connected to the descriptions of revisits to the museum and the associated expectations but also unfortunately events occurred during the visit or in connection to the visit itself, such as museum's disorganization in supporting visitors, lack of information or encounters with rude personnel.
- *Museum Services (3 latent topics)*: museum reviews address with an average probability 23% aspects connected to services offered by museums, such as ticketing, guided tours, accessibility and transports.

Table 5 "Bottom-up" quality dimensions derived from the Italian text of online reviewers of museums. Italian stem words and reviews excerpts have been translated to English to increase readability and comprehension, but the algorithm elaborated Italian texts.

| "Bottom-up" Quality Dimension Name | Most probable words (English translation) | Latent Topics Names | Excerpts of reviews |
|---|---|---|---|
| *Museum Cultural Heritage* | museum, exposition, floor, room, collection, church, century, fresco, building, structure, marvel, wonder, beauty, artwork, art, gallery, masterpiece, castle seen, inner, landscape, panorama suggest, external, garden, villa, park, palace, fountain | **1.** Artistic Collection<br>**2.** Exhibitions & Findings<br>**3.** Castles & Views<br>**4.** Churches & Religious Antiquity<br>**5.** The Museum from Outside<br>**6.** Museum's History & Tradition | "Castle built around 1850-60 in a medieval style wanted by the then Habsburg Empire, more precisely at the behest of Maximilian of Habsburg Archduke of Austria ..." (probability of observing *Museum Cultural Heritage* in review: 64%)<br>"… The museum presents a great variety of works ranging from painting, sculpture and architecture. In addition, you can admire some of the most famous masterpieces of the art world such as ... " (probability of observing *Museum Cultural Heritage* in review: 59%)<br>"... the temple was rebuilt in the form in which we can admire it today by the emperor Hadrian (128 AD) under whose reign the empire of Rome reached the height of its splendor ..." (probability of observing *Museum Cultural Heritage* in review: 69%) |
| *Personal Experience* | visit, day, suggest, time, place, emotion, experience, see, years, remain, pity, just, personal, turist, unfortunately | **7.** Emotional Visits<br>**8.** Revisits & Expectations<br>**9.** *At least once!*<br>**10.** *Unfortunately* | "... And every time, looking up at its internal vault, I am amazed at how a closed place can convey that sense of immense space and deep breath to me." (probability of observing *Personal Experience* in review: 44%)<br>"I remembered seeing a marvel in a state of abandonment, this year I revisited it and I was amazed ..." (probability of observing *Personal Experience* in review: 41%)<br>"… Reviewing the excavations is a great thrill every time, despite the poor signage and maps that are not always clear. However, the organization turned out to be even worse ... " (probability of observing *Personal Experience* in review: 47%) |
| *Museum Services* | see, guide, appreciate, organize, path, accompany, explain, tour, ticket, entry, euro, queue, reservation, cost, proce, entry, site, park, reach, close, foot, walk, convenient, easy | **11.** Accessibility & Transports<br>**12.** Guided Tours<br>**13.** Ticketing (purchase, price, book) | "... you can have free and privileged access to the cash desks ... the site is easily accessible from Naples by bus ..." (probability of observing *Museum Services* in review: 46%)<br>"... our guide was extremely engaging, able to actualize what we saw ..."(probability of observing *Museum Services* in review: 40%)<br>"… Having arrived 1 hour earlier we got in line anyway! Once the tickets have been taken, we are told the entrance ... " (probability of observing *"Museum' Services"* in review: 51%) |

The identification of these three "bottom-up" dimensions from the own words of museum reviews shows that museum visitors emphasize various aspects of the experience, moving beyond the services offered by museums delineated by the "top-down" categories. Specifically, the "bottom-up" analysis reveals that museum reviewers also consider cultural heritage aspects and personal experiences when evaluating the quality of the museum experience. This implies that, according to the visitors' perspectives, the museums' quality dimensions are not only limited to museum services, but extend to the museum cultural heritage and to the personal experience felt during the visit. This already offers interesting results for policy makers and museum experts, who should consider the existence of all the three "bottom-up" dimensions, rather than limiting the attention to museum services.

This suggestion is strengthened by the observation of the results of the "bottom-up" analysis (Table 6), which show that a review of museums discusses on average more about museum cultural heritage aspects (46% average probability) and personal experiences (31% average probability) rather than museum services (23% average probability). These results are interesting for museum experts and policy makers since they show an average predominance of emotional and heritage aspects of the visit experience on the services provided by museums within the own words of online reviewers. Moreover, these "bottom-up" results underline the bias of museum experts and policy makers when considering museum quality dimensions, aspect that will be the focus of the following section.

Table 6 Summary statistics of the probability distribution of the three "bottom-up" dimensions of museum quality, obtained from the sum of the per-document topic probability distributions of each topic associated to the corresponding "bottom-up" dimension.

| "Bottom-up" Quality Dimension Name | Min | Q1 | Q2 | Mean | Q3 | Max |
|---|---|---|---|---|---|---|
| *Museum Cultural Heritage* | 15.31% | 41.83% | 46.15% | 46.12% | 50.12% | 83.54% |
| *Personal Experience* | 9.37% | 27.31% | 30.55% | 30.80% | 33.94% | 58.46% |
| *Museum Services* | 6.47% | 19.59% | 22.05% | 23.09% | 25.61% | 59.46% |

### 3.3 RQ 3. To what extent do museum quality dimensions evaluated from online reviews in a "bottom-up" fashion differ from those identified in a "top-down" fashion?

The "top-down" and "bottom-up" approaches show contrasting results: while policy makers define museum quality dimensions in terms of services offered by the museum, visitors are on average more focused on emotional and heritage aspects of the visit. The application of the "top-down" approach started from a set of keywords attributable to museum service and defined from the standards issued by the policy maker; this supervised approach resulted into the 63% of online reviews (*Other aspects* category) that did not follow into any of the predefined categories. The "bottom-up" approach overcomes this limitation searching directly from the own words of visitors and relying on a probability distribution. Without even acknowledging how many or which could be the dimensions of quality of a museum, users emphasize various aspects of their visit, which hinder the evaluation of the perceived quality of museums. These hidden perspectives captured through LDA show that a museum review discusses on average more about museum cultural heritage aspects (46% average probability) and about personal experiences (31% average probability) than referring to services offered by the museum (23% average probability).

To further understand the differences between the two approaches not only in terms of models (see also Section 2.3) but also in terms of results, we analysed the museum quality dimensions emerging in a "bottom-up" fashion from the visitors' perspective and focused our attention on the reviews classified in a "top-down" fashion as *Other Aspects* (Figure 6). Looking at these reviews in a "top-down" logic, the policy maker would not have been able neither to detect any of the service classes nor to grasp the aspects of actual interest of museum visitors. On the contrary, using a "bottom-up" approach policy makers would have observed that the most discussed aspects are connected to the heritage of the museum (48% average probability of observing *Museum Cultural Heritage*) and to the personal experiences felt during the visit (31% average probability of observing *Personal Experience*), while the attention to museum services is just limited (21% average probability of observing *Museum Services*). Beyond this, the "bottom-up" approach here proposed also offers policy makers the opportunity to further explore the hidden aspects discussed by online reviewers of museums within each of the "bottom-up" dimensions. For example, the reviews classified as *Other aspects* reveal to address with high probability latent aspects connected to the museum's history (8.5% average probability of observing latent topic *Museum's History & Tradition*) and to artworks (8.1% average probability of observing latent topic *Artistic Collection*), but also to refer frequently to the emotions felt during the cultural experience (8.5% average probability of observing latent topic *Emotional Visits*). Since museum reviewers also recognize *Museum Services* (21% average probability of observing the dimension) among the "bottom-up" museum quality dimensions, the policy maker is also able to detect specific latent aspects connected to this dimension. These analyses reveal an average probability of observing the latent topic *Accessibility & Transports* equal to 7.4%, *Guided Tours* equal to 7.2% and *Ticketing (purchase, price, book) to* 6.6%. The following are examples (translated in English for clarity) of Italian reviews classified through the "top-down" approach as addressing *Other Aspects* but which show a high probability of discussing the "bottom-up" dimension of *Museum Services*.

*"From the Porta San Paolo railway station (from the Piramide metro stop) take the train to Ostia Antica, after a journey of about half an hour. In ancient times it was the ancient port of Rome and in it the goods flowed and passed to and from the whole empire. The ruins are well preserved and all the activities of the time are recognizable from them, from the storage warehouses to the bathrooms. public buildings, amphitheaters, fire stations, port corporations. I leave the rest to your curiosity, I bet you will be charmed"* (30.7% probability of observing latent topic *Accessibility & Transports*)

*"Nice initiative by the students of the Rodolico scientific high school. We were welcomed with kindness and cordiality by the students, appreciating their competence."* (20.1% probability of observing latent topic *Guided Tours*)

*"admission 9 euros per person and 4 euros for children seems a bit excessive to me .. to possibly add 1 euro for transport by bus because if you proceed on foot the path to take is not at all simple and in good shape. strollers is impossible!!!"* (19.2% probability of observing latent topic *Ticketing (purchase, price, book)*)

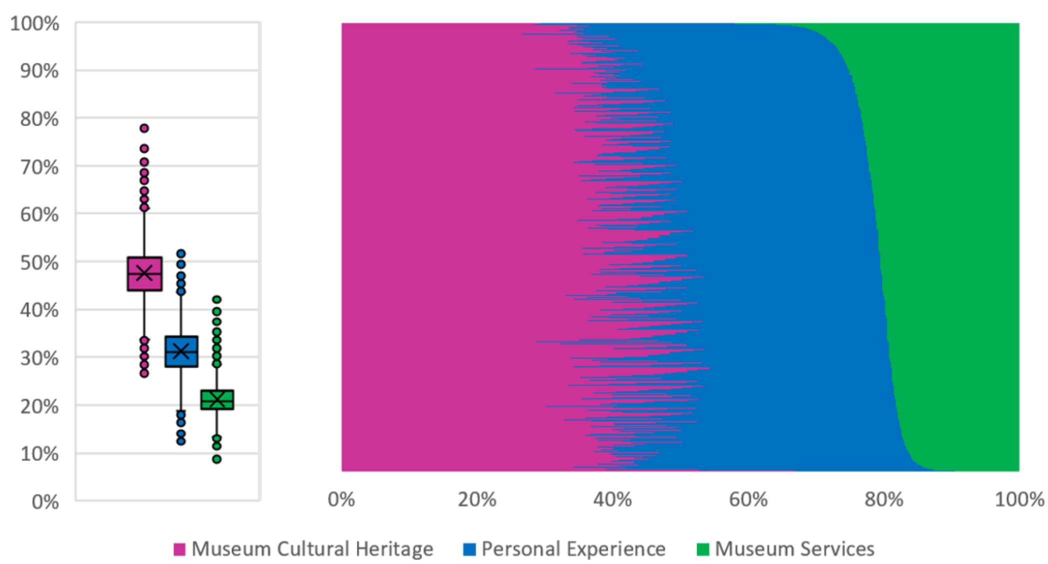


Figure 6 Comparison of the distributions of the "bottom-up" museum quality dimensions over reviews classified as addressing Other Aspects according to the "top-down" approach.

## 4 CONCLUSION

Although automatic models for text analytics have proven to be precious in exploring quality dimensions in various applicative settings (e.g. [Bi et al., 2020]; [Galati and Galati, 2019]), to the best of our knowledge there is still a lack of studies that systematically analyses online reviews for the automatic evaluation of quality dimensions of museums comparing the expectations of policy makers and the perceptions of museum visitors. This study addresses this gap, investigating museum quality dimensions among the online reviews of 100 Italian museums over a time-period of one year (2019).

The implementation of an automatic system to collect data from TripAdvisor and to enrich the collected information with the language of review and with sentiment scores, results in a dataset of 14,250 online reviews of museums for which the language has automatically been recognized to be Italian, enriched with a sentiment score provided by a BERT model specifically trained for Italian language. Inspecting data shows that Italian museums are positively evaluated by the public, as reviews show on average both high ratings (4.42 stars over 5) and high sentiment scores (0.7499 on a range from -1 to +1).

To replicate the "top-down" and "bottom-up" approach for the identification of museum quality dimensions, we selected respectively (i) a supervised keyword-based non-overlapping multiclass classifier for text of reviews, based on a set of keywords defined by the policy maker and (ii) an unsupervised statistical topic model for the automatic modelling of latent themes of discussions within the online user-generated text of reviews.

The results of the supervised analysis already underlined a risk for policy makers when adopting the "top-down" approach: searching for specific keywords in the review text, policy makers find just a limited amount of reviews associated to the desired service-related classes, while the majority of reviews address aspects that go beyond the specifications of the policy makers. This already evidences that visitors are not only interested in service-related aspects like ticketing, communication or activities offered by the museum, but are also interested in other aspects. Indeed, the "top-down" and "bottom-up" approaches to the analyses of reviews show contrasting results: while policy makers define museum quality dimensions connected just to services offered by the museum, visitors are on average more focused on emotional and heritage aspects of the visit.

To further explore the perspective of online reviewers and the misalignment with respect to the policy maker's perspective, we investigated in an automated "bottom-up" fashion the museum quality dimensions hidden within the content of online reviews. Without imposing any predefined category, we automatically detected the museum quality dimensions in a "bottom-up" fashion from the online voice of reviewers. Specifically, by choosing the LDA procedure to define a set of topic-based quality dimensions of museums, we simulated the evaluation of quality dimensions of museums pursued by museum visitors. Indeed, without even acknowledging how many or which could be the dimensions of quality of a museum, users evaluate their visit emphasizing various aspects, which hinder the perceived quality of museums. The results of this "bottom-up" approach show that a museum review discusses on average more about museum cultural heritage aspects (46% average probability) and about personal experiences (31% average probability) than referring to services offered by the museum (23% average probability). Moreover, the "bottom-up" approach here proposed also offers the opportunity to policy makers to explore the hidden aspects underlined by the online reviewers of museums within the "bottom-up" dimensions. This allows for instance to identify among reviews addressing *Other Aspects* high interest towards museum history and artworks (8.5% and 8.1% average probability of observing latent topic *Museum's History & Tradition* and *Artistic Collection* respectively), emotions felt during the cultural experience (8.5% average probability of observing latent topic *Emotional Visits*) and attention towards specific Museum Services, such as *Accessibility & Transports* (7.4% average probability), *Guided Tours* (7.2%) and *Ticketing (purchase, price, book)* (6.6%).

These findings underline the limited information policy makers can grasp from the online public if they just evaluate quality dimensions using "top-down" strategies instead of "bottom-up" approaches: with a "top-down" logic, policy makers only see that visitors are not highly interested in museum services, but they do not collect additional information on which are the relevant aspects according visitors. Instead, listening to the voice of the public of museums in a "bottom-up" logic offers policy makers the opportunity to understand not only which are the dimensions of interest of visitors but also allows them to define the visitors' priority towards these dimensions: the analysis reveal that policy makers should consider also emotional and heritage aspects of the visit experience, which are on average the most recurring aspects of discussion among online reviewers of museums.

Our research contributes to the discussions on the impact that different data analytics approaches have in supporting organizations' decision making processes. Our work shows how online data and their automatic elaboration could enable organizations to enrich their decision-making processes considering also the perspective of the general public. Showing policy makers and managers a way of considering the voice of the public opinion, our work aims at creating a bridge between technical and managerial aspects associated with the adoption of online data processing.

From a scientific perspective, the contribution of our work is twofold: (1) to demonstrate that online reviews can actually provide valuable insights for the evaluation of service quality dimensions defined by decision makers; and (2) to show that a bottom-up approach starting directly from textual expressions in reviews is able to identify further dimensions and quality aspects, that go beyond the typical service-centred analysis performed by businesses and institutions.

Considering the specific setting of museums, our research shows a particular use case of technologies and social media information to support decision-makers and managers in shaping their decisions. Indeed, our research shows decision-makers that an aware exploitation of online user-generated data could offer insightful results on the perspective of users with respect to quality dimensions. Specifically, moving beyond manual analysis of online reviews of museums (e.g., [Su and Teng, 2019]), the automatic analysis of reviews we presented allows decision makers to: (1) widen the understanding of the public opinion around the quality dimensions of institutions they manage, suggesting a preliminary identification of museum quality category also for museums, as already happens in other sectors (e.g. [Bi et al., 2020]; [Galati and Galati, 2019]); (2) record the (mis)alignment between the expected quality dimensions – defined using a top-down approach – and the quality dimension perceived by the users of institutions – derived from using a bottom-up approach – and therefore underlines the importance of exploiting the voice of online users since it can reveal unexpected aspects.

**ACKNOWLEDGMENTS**

The authors would like to thank the Italian Ministry of Cultural Heritage and Activities and Tourism for the collaboration shown along the design of the supervised classification described along this study.

**REFERENCES**

Arun, R., Suresh, V., Veni Madhavan, C. E. & Narasimha Murthy, M. N. (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. In Advances in knowledge discovery and data mining, Mohammed J. Zaki, Jeffrey Xu Yu, Balaraman Ravindran and Vikram Pudi (eds.). Springer Berlin Heidelberg, 391–402. http://doi.org/10.1007/978-3-642-13657-3_43

Bi J.-W., Liu Y., Fan Z.-P., Zhang J. (2020) Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry. Tourism Management. DOI: L10.1016/j.tourman.2019.104006

Bi, J.W., Liu, Y., Fan, P.Z., and Zhang, J. (2019), "Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews"; Tourism Management, 70, pp. 460-478

Blei, D. (2012). Probabilistic topic models. Communications of the ACM, 55(4):77–84.

Craig Wight, A. (2020). Visitor perceptions of European Holocaust Heritage: A social media analysis. Tourism Management, 81(May). https://doi.org/10.1016/j.tourman.2020.104142

Deveaud, R., SanJuan, É. & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. Document numérique 17, 1: 61–84. http://doi.org/10.3166/dn.17.1.61-84

Devlin, Jacob and Chang, Ming-Wei and Lee, Kenton and Toutanova, Kristina (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv preprint arXiv:1810.04805Eleni, M., and Costantine, L. (2013), "Factors affecting museum visitors' satisfaction: The case of greek muse-ums", Tourismos, 8(2), pp. 271-287

Ferguson, M., Pichè, J., and Walby, K. (2015), "Bridging or fostering social distance? An analysis of penal specta-tor comments on Canadian penal history museums", Crime, Media, Culture: An International Journal, 11(3), 357-374

Galati F., Galati R. (2019) Cross-country analysis of perception and emphasis of hotel attributes. Tourism Management

Giglio, S., Bertacchini, F., Bilotta, E., and Pantano, P. (2019), "Using social media to identify tourism attrac-tiveness in six Italian cities, Tourism Management", 72, pp. 306-312

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences 101, suppl 1: 5228–5235. http://doi.org/10.1073/pnas.0307752101

Guizzardi, A. and Mazzocchi, M. (2010). Tourism demand for Italy and the business cycle. Tourism Management. 31(3), 367-377

Hou, Z., Cui, F., Meng Y., Lian T., Yu C. (2019) Opinion mining from online travel reviews: A comparative analysis of Chinese major OTAs using semantic association analysis, Tourism Management, DOI: https://doi.org/10.1016/j.tourman.2019.03.009

Jia, S. (2020), "Motivation and satisfaction of Chinese and U.S. tourists in restaurants: A cross-cultural text mining of online reviews", Tourism Management, pp.1-20, https://doi.org/10.1016/j.tourman.2019.104071

Juan, C., Tian, X., Jintao, L., Yongdong, Z., & Sheng, T. (2009). A density-based method for adaptive lda model selection. Neurocomputing — 16th European Symposium on Artificial Neural Networks 2008 72, 7–9: 1775–1781. http://doi.org/10.1016/j.neucom.2008.06.011

Kirilenko, A. P., Stepchenkova, S. O., & Dai, X. (2021). Automated topic modeling of tourist reviews: Does the Anna Karenina principle apply? Tourism Management, 83(September 2020). https://doi.org/10.1016/j.tourman.2020.104241

Liu Y, Teichert T., Rossi M., Li H., Hu F., (2017) Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews, Tourism Management, DOI: http://dx.doi.org/10.1016/j.tourman.2016.08.012

Moreno Gil, S., Brent-Ritchie, J.R., and Almeida-Santana, A. (2019), “Museum tourism in Canary Islands: assessing image perception of Directors and Visitors”, Museum management and curatorship, 34(5), pp. 501-520

Nowacki, M., & Kruczek, Z. (2021). Experience marketing at Polish museums and visitor attractions: the co-creation of visitor experiences, emotions and satisfaction. Museum Management and Curatorship, 36(1). https://doi.org/10.1080/09647775.2020.1730228

Oren, G., Shani, A., & Poria, Y. (2021). Dialectical emotions in a dark heritage site: A study at the Auschwitz Death Camp. Tourism Management, 82(May 2020). https://doi.org/10.1016/j.tourman.2020.104194

Richards, G., King, B., & Yeung, E. (2020). Experiencing culture in attractions, events and tour settings. Tourism Management, 79(January 2019), 104104. https://doi.org/10.1016/j.tourman.2020.104104

Schuckert, M., Liu, X, and Law, R. (2015), “Hospitality and Tourism Online Reviews: Recent Trends and Future Directions”, Journal of Travel & Tourism Marketing, 32(5), pp. 608-621

Sheng, Chieh-Wen & Chen, Ming-Chia. (2012). A study of experience expectations of museum visitors. Tourism Management. 33. 53–60. 10.1016/j.tourman.2011.01.023.

Simeon, M. I., Buonincontri, P., Cinquegrani, F., & Martone, A. (2017). Exploring tourists’ cultural experiences in Naples through online reviews. Journal of Hospitality and Tourism Technology, 8(2), 220–238. https://doi.org/10.1108/JHTT-10-2016-0067

Su Y., Teng W. (2018) Contemplating museums’ service failure: Extracting the service quality dimensions of museums from negative on-line reviews, Tourism Management, DOI: https://doi.org/10.1016/j.tourman.2018.06.020

Taecharungroj, V., & Mathayomchan, B. (2019). Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. Tourism Management, 75(April), 550–568. https://doi.org/10.1016/j.tourman.2019.06.020

Torre, A., and Scarborough, H. (2017), “Reconsidering the estimation of the economic impact of cultural tourism”, Tourism Management, 59, pp. 621-629