TEMA
Technologies Engineering Materials Architecture

Firenze University Press
https://oajournals.fupress.net/index.php/tema

Citation: D'Orazio, M., Di Giuseppe, E., & Muccioli, M. F. (2026). Mining Users' Perceptions Through Sentiment and Emotion Analysis to Address Heritage Conservation Strategies. *TEMA Technologies Engineering Materials Architecture* 12(1): 65-77. doi: 10.36253/tema-16921

Received: March 19, 2025

Accepted: April 18, 2025

Published: April 26, 2026

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

ORCID:
MDO: 0000-0003-3779-4361
EDG: 0000-0003-2073-1030
MFM: 0009-0009-2602-0126

Mining Users' Perceptions Through Sentiment and Emotion Analysis to Address Heritage Conservation Strategies

Marco D'Orazio, Elisa Di Giuseppe, Maria Francesca Muccioli*

Department of Civil and Building Engineering and Architecture (DICEA), Università Politecnica delle Marche, Italy
Email: m.dorazio@staff.univpm.it; e.digiuseppe@staff.univpm.it; m.muccioli@pm.univpm.it
*Corresponding author

Abstract. Monitoring architectural heritage is a crucial step in the planning of proper conservation strategies and resource allocation. Current protocols rely on periodic, though infrequent, expert-led inspections, which assess the state of conservation of heritage assets and inform intervention priorities. However, public perceptions, which may suggest alternative courses of action, are seldom considered. This study proposes an innovative methodology integrating public feedback into heritage monitoring via Natural Language Processing (NLP). The framework, applied to '70s heritage sites in Italy's Marche region, integrates Aspect-Based Sentiment Analysis (ABSA) and Aspect-Based Emotion Analysis (ABEA) to systematically analyze user-generated content, identifying heritage-related aspects and classifying sentiment (positive, negative, neutral) and emotions (e.g., joy, anger) from Google Maps reviews. Heritage-specific targets were first identified in user reviews using spaCy-based tokenization. Sentiment classification (positive, negative, neutral) was performed using a pre-trained Bidirectional Encoder Representations from Transformers (BERT) model, while emotions (joy, anger, sadness, fear) were identified using the FEEL-IT algorithm. User perceptions were effectively retrieved, revealing a generally positive sentiment and joy as the most dominant emotion. This approach enables large-scale monitoring based on continuously updated user feedback, which can be integrated into current monitoring protocols to adopt a more comprehensive decision-making approach.

Keywords: Architectural heritage, Heritage monitoring, Resources allocation, User sentiment, User emotion.

1. INTRODUCTION

Monitoring architectural heritage is essential to guarantee proper knowledge to inform decision-making and the allocation of financial resources for conservation interventions. Continuous surveillance allows for the early detection of potentially critical issues, enabling timely, targeted actions. Strategies that adopt a *preventive* rather than *curative* approach help avoid delayed restoration efforts, which often fail to address underlying issues and require substantial public funding [1].

Currently, monitoring protocols are managed by public administrations, which rely on experts to assess the state of conservation of heritage buildings through structured reports. Their evaluations constitute the base knowledge for stakeholders responsible for decision-making [2].

However, these documents often fail to reflect the actual status of the assets and become outdated as soon as they are compiled. This is not only due to the evolutionary nature of the built environment, but also to the infrequent site inspections by technicians.

Although seldom considered, the potential of strategies that involve end-users in monitoring and planning actions has been explored and supported by several studies and directives [3]. In fact, involving the general public not only fosters a democratic approach but also offers a more comprehensive perspective, highlighting overlooked issues or suggesting alternative courses of action for decision-making. Moreover, public engagement also provides an opportunity to monitor the outcomes of actions taken (or not taken) in heritage conservation and management.

Heritage architecture is a topic of widespread interest among the public, whose opinions are frequently reported on social media platforms in the form of short texts, comments, or reviews. The internet thus serves as a large and dynamic repository of unstructured data, which is continuously updated and provides near-realtime and unbiased information. Thanks to Artificial Intelligence (AI) technologies, and in particular to Natural Language Processing (NLP) techniques, it has now become possible to exploit this huge amount of user-generated content [4].

Increasingly popular applications of NLP are represented by Sentiment Analysis (SA) and Emotion Analysis (EA). SA, also called *Opinion Mining*, is the computational study of people's opinions, allowing for the detection of sentiment orientation (positive, negative or neutral) or sentiment intensity (e.g. sentiment polarization on a scale from 1, most negative, to 5, most positive). EA provides more detailed insights, detecting and classifying user emotions into distinct classes (e.g. joy, anger, sadness and fear, but they can vary depending on the adopted psychological model) [5-6].

Although widely used to assess user satisfaction in products and services, SA and EA have seen limited application in the construction sector, where they remain in their early stages.

For instance, D'Orazio et al. [7] employed different SA models to process user-generated reports from a computerized maintenance management system for a stock of university-administrated buildings and to define an intervention priority scale. A comparative analysis of the models' efficiency was performed, and the potential of such tools for this kind of tasks was confirmed. In a subsequent study, D'Orazio et al. [8] used a similar approach to prioritize maintenance requests, but included EA to categorize perception into joy, anger, fear and sadness categories, providing deeper insights.

In the context of heritage architecture, Rosin et al. [9] applied the SA model VADER (Valence Aware Dictionary for Sentiment Reasoning) [10] to online user reviews of the maritime museums network *Arca Adriatica*, examining how visitors' feedback could enhance management strategies. Mendes et al. [11] proposed instead the use of VADER to process Tripadvisor reviews of two Iberian monuments, identifying triggering factors of negative sentiment; the study revealed that maintenance actions restricting access to the site constitute the main cause of dissatisfaction among tourists.

SA can be performed at different levels. Ginzarly et al. [12] applied SA at the sentence level to the TripAdvisor reviews of two Monuments in Lebanon to categorize their perception as positive or negative and thus provide insights on cultural values and attributes expressed by users. Valdivia et al. [13] detected the sentiment polarity for three Spanish monuments by applying an Aspect-Based Sentiment Analysis (ABSA) on Tripadvisor reviews, identifying negative aspects and issues that required intervention by cultural managers.

While document level SA detects the overall sentiment for a document, and sentence level SA for a single phrase, ABSA aims at identifying what users exactly appreciate or not [5]. Within a document – or even a single sentence – multiple thoughts may actually be expressed (e.g. the sentence "The museum has fascinating exhibits, but the staff is unhelpful" conveys a positive sentiment for the aspect "museum", but a negative one for "staff"). Focusing the SA and EA on the selected *target* is crucial for obtaining reliable results. ABSA still represents a challenging field for research, due to the complexity of linguistic structures, such as sentences which present negations or which are factual but still imply opinions. Additionally, the lack of sufficient domain-specific training data constitutes further difficulties.

A proper revolution in the field came with the introduction of pre-trained models based on transformer technology: among them, it is recalled BERT (Bidirectional Encoder Representations from Transformers) language model [14]. Introduced by Google in 2017, BERT uses an encoder based on self-attention mechanisms and incorporates Masked Language Modelling (MLM) and Next Sentence Prediction (NSP), which enhanced the comprehension of word relationships and of the general context. Its versatility allows for fine-tuning across a range of tasks, including SA.

The advancements hereby presented offer just a brief glimpse into the potential of applying SA and EA tools to user-generated content related to architectural heritage, which currently lacks well-defined frameworks or methodologies [15].

This study aims to present a novel methodology which leverages pre-trained BERT-based models to perform ABSA and EA on user-generated reviews concerning architectural heritage. The results will provide an overview of user perceptions of the assets which are not influenced by non-architectural aspects (e.g. the function of buildings), and deeper insights will be given by the detection of conveyed emotions. The presented approach intends to offer a novel monitoring strategy which is more comprehensive and can complement current protocols and favor public engagement.

This paper is structured as follows: Sect. 2 describes methods, starting from the illustration of the proposed methodology (Sect. 2.1) and a description of the case study selected to test it (Sect. 2.2); Sect. 3 presents data resulting from the analysis along with its discussion; Sect. 4 concerns conclusions and future possible implications of the present research.

2. METHODS

2.1 Aspect-Based Sentiment and Emotion Analysis methodology

The proposed Aspect-Based procedure for SA and EA was tested on user-generated textual reviews from Google Maps related to heritage sites in Italy's Marche region. This dataset provides the basis for demonstrating the framework's applicability. Information regarding the selection of heritage sites and the collection of user-generated content will be described in greater detail in Section 2.2.

The methodology can be synthesized in three phases (Figure 1), which are briefly described here and will be detailed in the following sections:

1. *Heritage target detection* (Sect. 2.1.1): the NLP model spaCy was employed to detect words (tokens) within the text of reviews related to the heritage domain. These tokens represented the *targets* for the aspect-based analysis, enabling the recognition of segments of sentences to be processed by SA and EA algorithms.
2. *Sentiment Analysis* (Sect. 2.1.2): a pre-trained BERT model was employed for the SA of the previously identified segments of reviews' sentences. Sentiment scores of pieces of text were aggregated to obtain a single score for the analyzed heritage item.
3. *Emotion Analysis* (Sect. 2.1.3): the FEEL-IT [16] model was used to provide further insight on sentiment polarity and categorize segments into emotion classes (joy, anger, fear, sadness).

2.1.1 Heritage target detection

Two key steps characterize this phase: (i) identifying heritage-related tokens in the review texts, and (ii) isolating sentence segments containing these tokens.

To this aim, the use of the Python library spaCy was chosen [17]. SpaCy is an NLP library which supports more than 75 languages, including Italian, the language used in the analyzed reviews. One of its main features consists of breaking texts into basic units (tokens) and

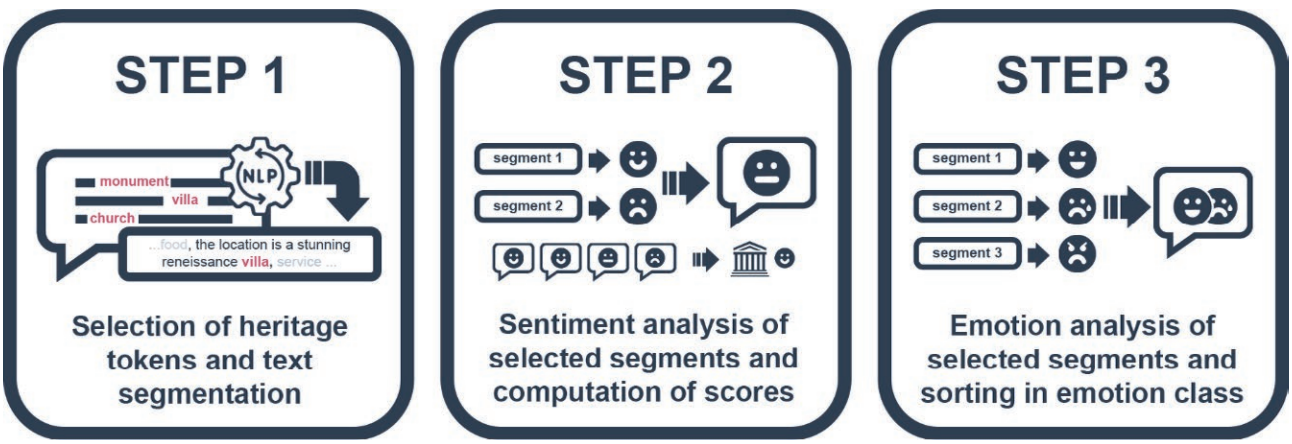

Figure 1. Representation of the three main phases at the base of the proposed methodology. © 2025, D'Orazio et al.

assigning a label (noun, adjective, etc.) to each through Part of Speech (POS) tagging.

By leveraging this function, words representing subjects or direct objects of sentences were detected, as they typically correspond to the parts of sentences most closely associated with opinions. The combination with the default Python module Collection allowed for the creation of an ordered list of terms, from the most to least frequent. The final selection of items followed these criteria: (i) architectural heritage related words or its physical surroundings (e.g. "gardens"); (ii) terms referring to immovable works of art or historically significant furnishings that contribute to the site's value (e.g. "sculptures", "altars"); (iii) terms associated with user experiencing the sites (e.g. "visit").

The identified terms served as targets to identify the parts of the review sentences to be isolated for later analysis. The boundaries of these segments were defined by the presence of punctuation marks or conjunctions (POS tagging identified), as these usually represent the borders of completed thoughts.

This passage allowed isolating parts of reviews related to heritage topics in the text (Table 1).

2.1.2 Sentiment Analysis

The SA processing of the previously identified sentence segments (Sect. 2.2.1) was performed using the pre-trained model BERT-base-multilingual-uncased-sentiment [18] from the Hugging Face platform. As the name itself suggests, the main features of the model are its capability to support multiple languages, including Italian, and to be unaffected by the use of uppercase or lowercase characters. The main reason that favored its selection was the provision of sentiment scores on a 5-star scale, analogous to the star ratings directly assigned by users, useful for a later comparison between the detected and the expressed perceptions of the users about the assets.

After analyzing each segment in a single review, their mean was calculated to obtain a single score for all of them. Once all the review scores had been computed

Table 1. Example of words (subject and direct objects) retrieved by spaCy and example of the selection of terms related to heritage and the associated sentence segments for the heritage site Church of Portone, in Senigallia. © 2025, D'Orazio et al.

Heritage Architecture	List of all detected tokens (subject and direct object) with frequency (within brackets) – English translation	List of tokens selected for the individuation of segments – English translation	User	Example of review – English translation
Church of Portone	parish priest (3); church (2); atmosphere (2); title (1); function (1); consideration (1); celebration (1); parish (1); theatre (1); lectures (1); vision (1); punishment (1); share (1); chapel (1); tabernacle (1); architecture (1); parking (1); works (1); youth (1); favourite (1); confessor (1); peace (1); name (1); enchanting (1)	<u>Building related:</u> church; chapel; architecture <u>Works of art/furniture:</u> tabernacle; works <u>Users experience:</u> atmosphere; sorrow; favourite; enchanting	User_01	Nice place. The church is nice and cosy. There is an atmosphere of recollection just entering it. It is a reference point for the community. The parish is very active in social work. The theatre next door often hosts themed conferences. I recommend watching them. It is well worth it.
				Selected segments for analysis
				"The church is nice and cosy"; "There is an atmosphere of recollection just entering it"; "It is well worth it"
				Segments excluded from analysis
				"Nice place"; "It is a reference point for the community"; "The parish is very active in social work"; "The theatre next door often hosts themed conferences"; "I recommend watching them"
			User_02	7.15 p.m. Sunday Mass. Full of people, priest engaging and smart. Singing few and well done. The architecture is simple. Outside there lots of parking spaces.
				Selected segments for analysis
				"The architecture is simple"
				Segments excluded from analysis
				"7.15 p.m. Sunday Mass"; "Full of people"; "priest engaging"; "smart"; "Singing few"; "well done"; "Outside there lots of parking spaces"

Figure 2. Calculation of a single sentiment score for an asset. © 2025, D'Orazio et al.

in this way, the sentiment score for the heritage item was obtained by averaging the scores of the corresponding reviews (Figure 2).

To categorize reviews into sentiment polarity classes, scores 1 and 2 were considered negative, 3 neutral, and 4 and 5 positive.

2.1.3 Emotion Analysis

Parallel to SA, the segments were processed with the model FEEL-IT to classify the emotions conveyed in the texts. FEEL-IT is a fine-tuned UmBERTo model (itself a fine-tuned version of BERT for Italian language) trained on tweets annotated with labels corresponding to the four basic emotions: joy, anger, fear and sadness [19].

Each of the segments identified in the first phase (Sect. 2.2.1) was processed using the selected algorithm. Just as a single sentence can convey both positive and negative sentiments, it can also express multiple emotions simultaneously. Since emotions are qualitative rather than quantitative, it was not possible to apply a procedure similar to the one used for calculating sentiment scores to derive a single result value (Sect. 2.2.2). Handling multi-opinionated sentences remains a significant challenge in this context [6, 20].

Therefore, the results of EA are presented as computed, with multiple emotion labels for each review.

2.2 Case study

The selected case study on which the proposed methodology was tested consists of catalogued heritage architecture located in the region of Marche in Italy. Italy is one of the richest countries in terms of heritage, and the region Marche alone counts more than 1000 monuments, 106 castles, 15 fortresses, thousands of churches and 72 theatres, as well as presenting the highest museums-population ratio and UNESCO sites [21].

The Italian authority responsible for heritage protection is the Ministry of Culture (MiC), which operates through peripheral offices, such as territorial Superintendencies for Archaeology, Fine Arts, and Landscape (SABAP, *Soprintendenze Archeologia, Belle Arti e Paesaggio*), in charge of the conservation and promotion of heritage sites.

These cultural stakeholders base their activities on data provided by platforms which collect information on protected sites. The heritage items selected for this study were sourced from the General Catalogue of Cultural Heritage [22]. This platform is managed by the MiC's Central Institute for the Catalogue and Documentation (ICCD) and organizes data collected on structured sheets during periodic on-site cataloguing campaigns conducted by local authorities (superintendencies, provinces, etc.), which include information on the state of conservation of assets [23].

The complete dataset for Marche's architectural heritage from the General Catalogue, comprising 6835 sites, was retrieved from the MiC open data platform at dati.beniculturali.it [24]. For each site, data such as the unique national ID, name, coordinates, state of conservation, notes on conservation status (e.g., causes of decay), and dates of compilation and updates were retained. This dataset represents official data collected by experts following official monitoring protocols.

Subsequently, a further selection was made from the official dataset to identify heritage items provided with user-generated content. The majority of the buildings listed in the catalogue are privately owned, precluding access to visitors and thus impeding discussions on social media.

Google Maps was chosen as the platform for retrieving user data. As a vast repository of location-specific information, it includes heritage sites and allows users to contribute content, such as five-star ratings, textual reviews, and even attached photos.

Matching items from the General Catalogue of Cultural Heritage and Google Maps could only be done by manual search. Automated procedures were hindered by the presence of inconsistencies in items' names (e.g. naming conventions as the use of "S." for "Saint" for churches and convents) or in coordinates data. Setting a tolerance distance to address the latter issue only made it worse, as the proximity of buildings in historic centers increased the risk of mismatches.

The final dataset of user-generated content comprised 70 items, including ratings, textual reviews, and usernames of the authors (hidden to protect their privacy), all scraped using the Google Chrome extension Instant Data Scraper. As most of the reviews were in Italian, the few reviews in other languages were extracted in their automatically translated form to maintain consistency within the dataset (multilingual SA is a possibility, but it is not the focus of this study).

The reduction from the 6835 items of the official catalogue dataset to the 70 of the user-generated content dataset is due to two main factors. First, the majority of the architectures listed in the official catalogue are private, preventing users from visiting them. Second, ministerial records also include *minor heritage* sites (it is noteworthy that Italian law protects public buildings over 70 years old), often less appealing to the general public and resulting in fewer reviews.

The results obtained by applying the methodology described in Section 2.1 to the case study dataset are presented in the following Section 3.

3. RESULTS

First considerations concerning the results obtained from the application of the proposed methodology for ABSA were conducted by comparing sentiment scores with the ratings directly assigned by users in the review process.

At first glance, the distribution patterns of the graphs are quite similar, confirming the methodology's effectiveness in capturing user perceptions, which is, for the most part, positive. However, some differences remain. The ABSA score histograms show fewer items in the most extreme categories (1- and 5-star ratings), with a higher concentration in the middle range. This suggests that ABSA tends to present less *extreme* perceptions, likely because the analysis focused on heritage-related aspects, excluding other factors that might have led to more polarized opinions.

Further analysis incorporated these results alongside findings from the Aspect-Based Emotion Analysis (ABEA). While, as noted in Sect. 2.2.3, it is not possible to force multi-emotion results into a single category, the most common emotion per heritage asset was retrieved to provide a qualitative overview of ABEA. Joy emerged as the dominant emotion, far surpassing the others. Given that joy is universally recognized as positive, while anger and fear could be considered negative emotions [19], ABEA also proved effective in capturing user perceptions. It is important to note that sadness does not always represent a negative emotion. For instance, in this case study, some reviews expressing nostalgia were categorized as sadness. These reviews did not convey negativity but rather a sense of affection and longing for something lost over time. Nevertheless, joy remains the most frequently detected emotion, far surpassing sadness in the ABEA results.

All these considerations are based on the data presented in Figure 3.

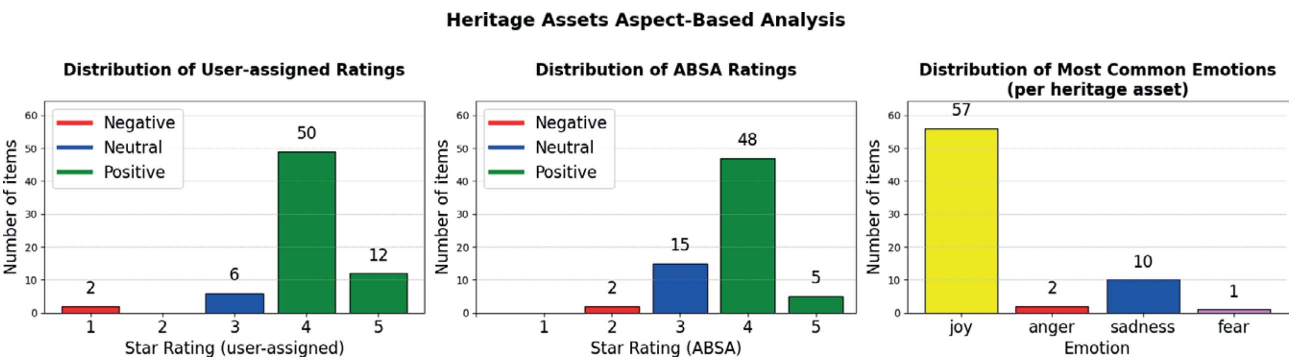

Figure 3. Histograms representing the distribution of ratings directly assigned by users (left), of resulting scores of the ABSA (center) and most common emotions for heritage assets (right). To sort results into sentiment polarity classes 1- and 2-star ratings were considered as negative, 3 as neutral and 4- and 5-star as positive. © 2025, D'Orazio et al.

Deeper insights are provided by dividing in categories the heritage architecture dataset based on the destination of use of buildings. The categories include directional buildings (4 items), fortifications (5), schools (4), monuments (1), museums (4), productive buildings (1), churches and monasteries (39), receptive structures (8), public buildings (1), theatres (2), and historic villas (1). Again, the distribution of user-assigned ratings (only the ones associated with the processed textual reviews), ABSA scores and ABEA results were analyzed in parallel. By making this distinction, it was easier to highlight cases worthy of attention.

One of the most notable cases is represented by the Post Office building of Ancona, the only one falling under the public building category. This item received the most negative evaluations by users, averaging just 1 star, which increased only to a 2-star score after ABSA processing of reviews (Figure 4). This shift may be attributed to the aspect-based approach; however further examination of the texts revealed that even when targeting the analysis through the use of heritage-related tokens (e.g. "place", "war"), users primarily expressed complaints about the service provided at the office (Table 2). This is further supported by the ABEA analysis of the reviews, which clearly indicates a predominance of anger in the examined texts (Figure 4).

Another interesting example is provided by the former Mancini furnace in Pesaro (productive buildings category). This site received only two reviews, averaging a rating of 4, as shown in Table 3. The first review (User_01) is paired with a highly positive rating from the author but was assessed as neutral by the ABSA, likely because the text is mainly descriptive, with the user implicitly expressing an appreciation for the allure of ruins. ABEA offers additional interpretations: the review conveys sadness in relation to the site's abandonment, anger over the decay caused by plant overgrowth, and fear regarding the dangers involved in accessing an opening to enter. This last emotion does not directly pertain to the asset's conservation status but rather to a dangerous action involving the site. While not strictly linked to conservation, such observations can provide valuable insights into how people experience a heritage asset in its current state of decay. The second review (User_02) shows a close alignment between the user's rating and the ABSA score, while the ABEA interpretation reflects sadness over the structure's lack of redevelopment.

Figure 5. provides a graphical representation of the data discussed above.

Final assessments compared the results of the proposed methodology for ABSA regarding user perceptions with expert evaluations on the state of conservation, as recorded in the official catalogue.

To achieve this, terms used in the ICCD sheets were categorized as positive, negative, or neutral, consistent with the ABSA results. As shown in Table 4, terms such as "good" and "excellent" were classified as positive, terms describing damage or explicitly negative conditions ("bad", "very bad") were classified as negative, and references to ongoing work or terms not fitting the other two categories were considered neutral. The term "ruin" was classified as negative, as it can be considered a synonym of "collapsed" from a conservation standpoint; however, it is worth noting that a ruin can have historic value and be in an optimal state of conservation [25].

Exploiting coordinate data, all items were mapped and marked with a color depending on the respective category (Figure 6). Results of the ABEA were excluded from this analysis due to the complexity of reporting multi-emotion-labelled data on a map.

The first observation is that the official record presents the most negative view of heritage conditions, particularly in the southern and inland areas of the region. By contrast, ABSA reflects a more positive outlook, with only two negative cases. These two cases, the Post Office of Ancona and the San Decenzio Cemetery in Pesaro, were given positive evaluations by experts but were further analyzed through user reviews. In the case of the Post Office, reviews primarily criticize the postal service, as discussed earlier in this section; for the cemetery, users reported various signs of structural decay (Table 5), which were further reflected by the predominance of anger and sadness resulting from the ABEA. Examination of the catalogue's compilation date revealed it to be from 2004, highlighting the outdated nature of the official records.

For sites marked by experts as in a negative state of conservation but classified as positive or neutral by ABSA, the conservation status notes in the official records were examined. It was found that the damages leading to these negative expert evaluations were largely caused by the 2016 earthquake that affected the southern part of the region. Although these records were more recent (averaging around 2018), they did not reflect recent restoration and refurbishment efforts that have since made these structures accessible to visitors.

4. CONCLUSIONS

Ensuring continuity in monitoring operations is crucial for the safeguarding of heritage and fosters preventive protocols. In addition to preserving cultural values,

Table 2. Example of reviews for the Post Office building of Ancona: the presented texts clearly show that, even when focusing the analysis on selected segments, these consist of complaints. © 2025, D’Orazio et al.

Heritage Architecture	Unique national ID	Selected tokens	User	User-assigned rating	ABSA score	Example of review – English translation	ABEA emotions
Post Office building	1100216467	post; war; place	User_01	1/5	5/5	You guys go to the sub-offices, here you wage war. Absurd	anger
			User_02	1/5	2/5	They don't answer the phone and the biggest shame is that no office in the yellow pages of the same building in Lagro XXIV Maggio has the working seriousness to pick up the phone and the smartest ones hang up the fax [...] Now work, people, there are those who are laid off and would like a job [in Italian “posto di lavoro”, workplace] like yours that would make this company much better, increasingly in a state of decay	sadness
			User_03	1/5	2/5	No one answers the phone, maybe the place is abandoned. Good service, that's love for one's work	sadness

Figure 4. Top: histograms representing the distribution of ratings directly assigned by users (left) and of resulting scores of the ABSA (right) for the Post Office building of Ancona; Bottom: Representation of the emotions associated with each of the analyzed reviews’ sentence segments for the Post Office building. © 2025, D’Orazio et al.

Table 3. Reviews for the former Mancini furnace of Pesaro. © 2025, D'Orazio et al.

Heritage Architecture	Unique national ID	Selected tokens	User	User-assigned rating	ABSA score	Example of review – English translation	ABEA emotions
Former Mancini furnace	1100221349	furnace; plants; brambles; ruins; opening	User_01	5/5	3/5	Together with xxx and xxx, I went to explore the disused Mancini furnace. Plants - unfailingly brambles - besieged the ruins. We do not give up and make our way through the thorns and mosquitoes. We reach an opening that is not too high, we cross it and we are inside	Anger, fear, sadness
						I went to explore the disused Mancini furnace	sadness
						Plants - unfailingly brambles - besieged the ruins	anger
						We reach an opening that is not too high	fear
			User_02	3/5	4/5	A ruin awaiting targeted redevelopment.	sadness

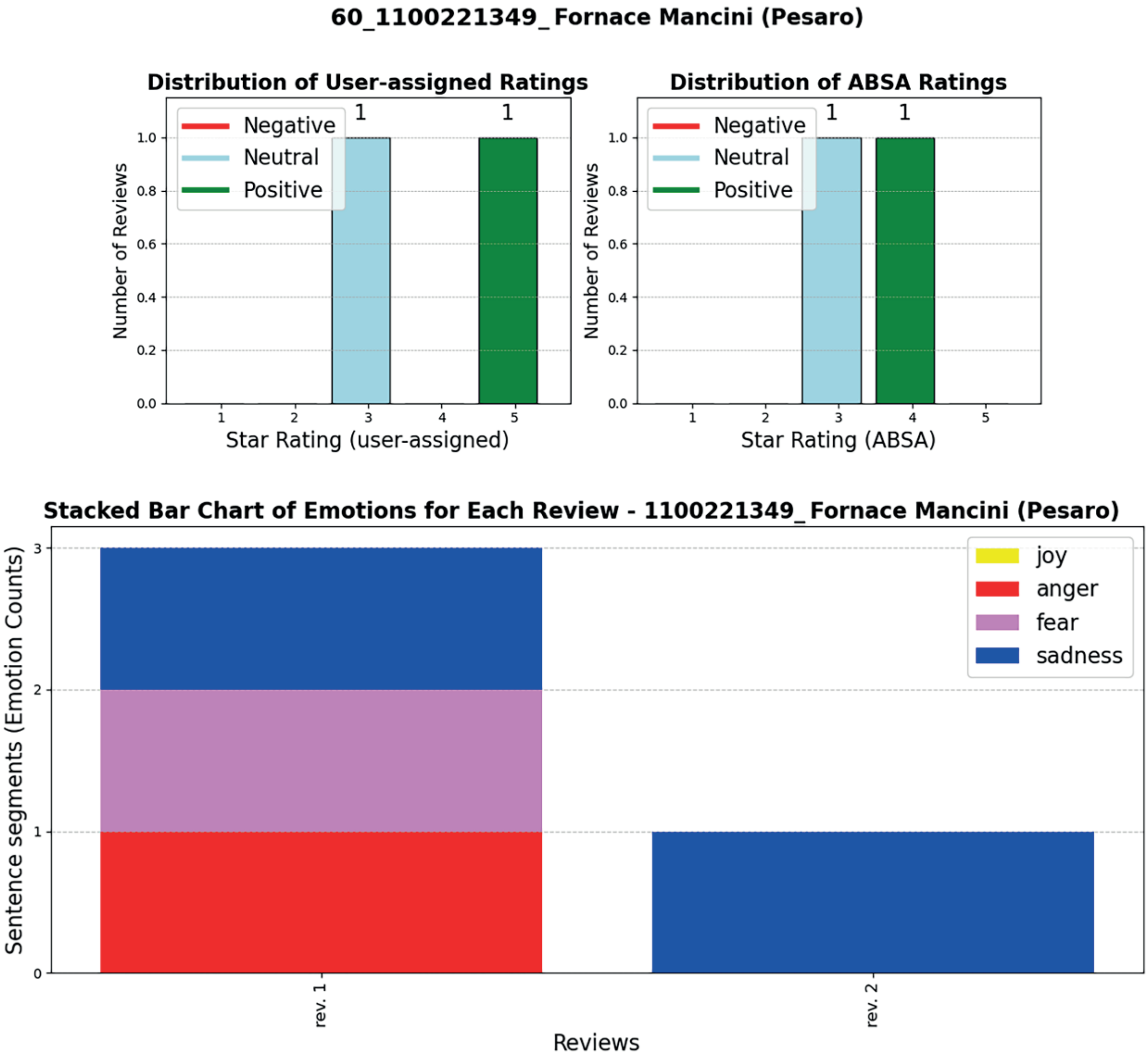

Figure 5. Top: Histograms representing the distribution of ratings directly assigned by users (left) and of resulting scores of the ABSA (right) for the former Mancini furnace in Pesaro; Bottom: Representation of the emotions associated with each of the analyzed reviews' sentence segments for the former Mancini furnace. © 2025, D'Orazio et al.

Table 4. Classification of conservation state terms from ICCD sheets. © 2025, D'Orazio et al.

Sentiment polarity	ICCD state of conservation	ABSA score
Positive	Excellent; Good	5 4
Neutral	Discrete; Absence of damage; Work in progress; Under renovation	3
Negative	Minor damage; Moderate damage; Medium damage; Damage of moderate level; Mediocre; Bad; Very bad; Severe damage; Very severe damage; Partly collapsed; Ruin; Collapse; Collapsed	2 1

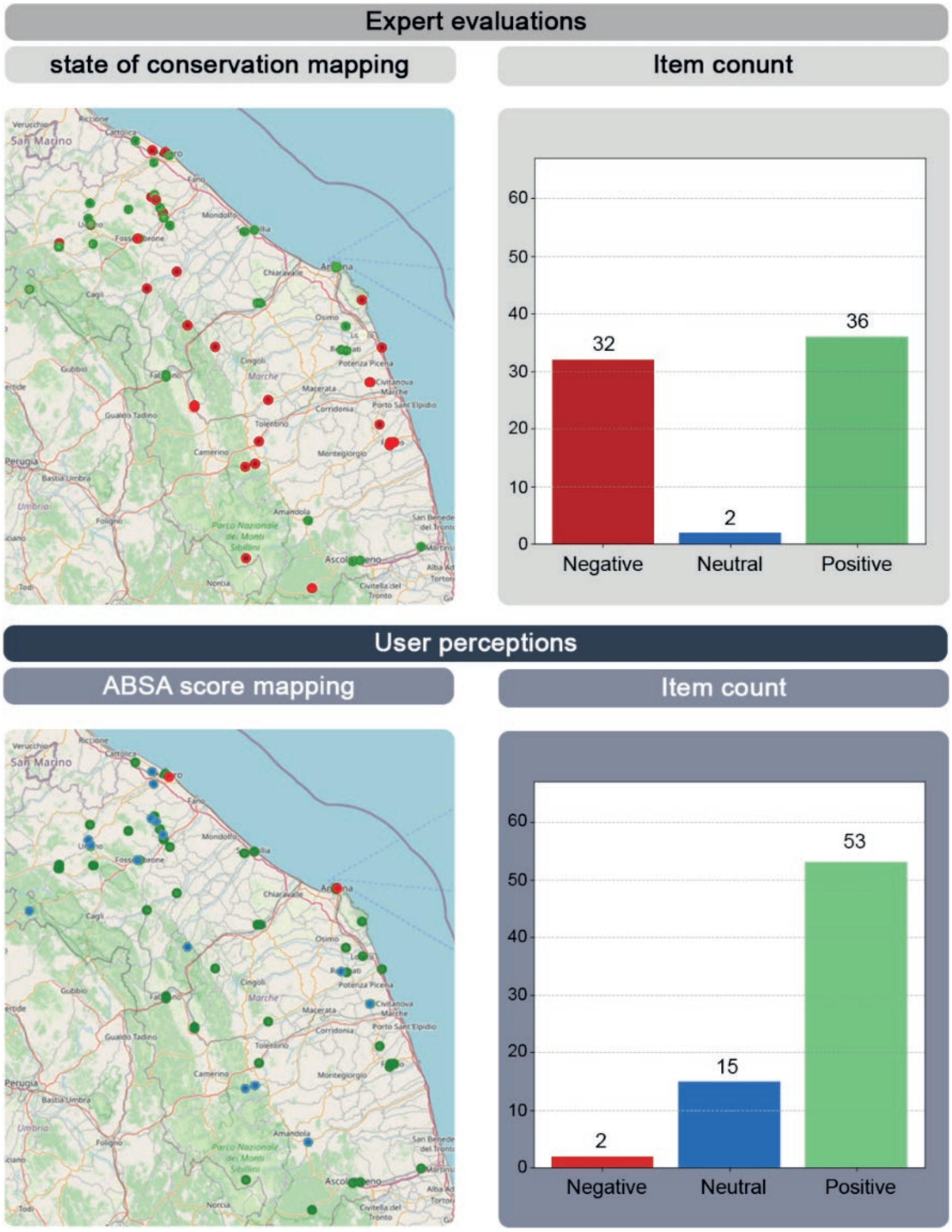

Figure 6. Mapping of expert and user data, respectively, concerning the evaluations on the state of conservation of heritage architectures and ABSA scores resulting from the proposed methodology. © 2025, D'Orazio et al.

Table 5. Example of reviews for the San Decenzio cemetery in Pesaro, reporting decay and damage to the structure. It is noteworthy that ABEA correctly interprets the use of the capital letters as a means of expressing anger. © 2025, D'Orazio et al.

Heritage Architecture	Unique national ID	Selected tokens	User	User-assigned rating	ABSA score	Example of review – English translation	ABEA emotions
San Decenzio Cemetery	1100221395	Part; cemetery; need; sinks; money; signs; night; exit; enters; pavilion; compressed	User_01	1/5	2/5	Cemetery in a pitiful condition, the new part needs constant work due to infiltrations, not to mention the old one taken by storm by pigeons that defecate on so many tombstones, in addition to the poor maintenance (see sinks always clogged)	sadness
						cemetery in a pitiful condition	sadness
						the new part needs constant work due to infiltration	sadness
						in addition to the poor maintenance (see sinks always clogged)	sadness
			User_02	1/5	2/5	IT IS IN INDECENT CONDITION: ESPECIALLY THE FLOWER PAVILION, THE LAST ONE BUILT, WHERE IT STILL RAINS INSIDE: they are the mirror of those who govern us, incapable of spending public money: all the last works have been resolved with compressed chipboard panels and painted with water-repellent paints that are already showing clear signs of subsidence but according to our politicians it is usable and well maintained	Sadness, fear, anger
						ESPECIALLY THE FLOWER PAVILION	anger
						incapable of spending public money	anger
						all the last works have been resolved with compressed chipboard panels	fear
						painted with water-repellent paints that are already showing clear signs of subsidence	sadness

timely actions and routine maintenance can prevent costly, delayed restoration efforts that may compromise assets or overlook key issues.

To support up-to-date and comprehensive monitoring protocols, a novel methodology was proposed to integrate user perceptions from social media platforms with official heritage conservation catalogues. This approach leverages SA and EA to process user-generated reviews from Google Maps, employing NLP techniques to isolate heritage-related topics from other aspects discussed in comments.

ABSA returned an overall positive perception of the heritage sites located in the Italian region of Marche, selected as a case study. With regard to the directly user-assigned 5-star ratings, results tended toward less *extreme* values, with fewer 1- and 5-star scores. Deeper examination of specific items yielded additional insights, such as the appeal of abandoned structures.

Comparing user perceptions with expert evaluations exposed outdated catalogue records, with notable discrepancies in resulting user perceptions. Some sites marked positively by experts met with user criticism over maintenance issues, alongside the conveyance of anger and sadness-related emotions; these catalogue entries dated back 20 years. Conversely, buildings damaged by the 2016 earthquake and marked negatively by experts received positive feedback from users, reflecting restoration efforts unrecorded in official documents.

This study presents a new methodology for ABSA and ABEA in heritage monitoring and identifies critical issues in current protocols. Future research may refine this model, developing automated matching procedures between official and user-generated datasets, a process currently hindered by issues related to naming conventions and coordinates. At present, this matching is done manually, making it time-consuming. Automating the selection of heritage tokens would also represent a significant improvement. By constructing and linking a heritage-related thesaurus, the selection of relevant tokens could form the foundation for a fully automated process. Once these issues are addressed and the model refined, the methodology could be applied to even larger datasets, such as at the national scale. Other potential implementations could involve exploring and complementing additional social media platforms, which could significantly enrich the insights derived from users' feedback and provide further validation.

This work lays the foundation for real-time monitoring platforms driven by user data, enhancing resource allocation and fostering participatory approaches by engaging the public, whose role currently remains underutilized in heritage management.

5. ACKNOWLEDGEMENTS

The present study has been supported by the Superintendence for Archaeology, Fine Arts and Landscape (SABAP, *Soprintendenza Archeologia, Belle Arti e Paesaggio*) for the Provinces of Ancona and Pesaro and Urbino. Authors extend their gratitude to the Superintendent Architect and Engineer Cecilia Carlorosi and to the referent of the Architectural Heritage Cataloguing Office Savino Petruzzelli.

6. FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

7. AUTHOR CONTRIBUTIONS

Marco D'Orazio: Conceptualization, Methodology, Formal analysis, Resources, Writing – review & editing, Supervision, Project Administration. *Elisa Di Giuseppe:* Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Project Administration. *Maria Francesca Muccioli:* Conceptualisation, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualisation.

8. REFERENCES

[1] Stovel H (2004) Monitoring world heritage. In: International Congress on World Heritage 2002: Shared Legacy, Common Responsibility, Venice, 2002. UNESCO World Heritage Centre and ICCROM, Venice

[2] Kioussi A, Karoglou M, Labropoulos K, Bakolas A, Moropoulou A (2013) Integrated documentation protocols enabling decision making in cultural heritage protection. Journal of Cultural Heritage 14(3):e141–e146. https://doi.org/10.1016/j.culher.2013.01.007

[3] Li J, Krishnamurthy S, Pereira Roders A, Van Wesemael P (2020) Community participation in cultural heritage management: A systematic literature review comparing Chinese and international practices. Cities 96:102476. https://doi.org/10.1016/j.cities.2019.102476

[4] Foroughi M, de Andrade B, Pereira Roders A (2023) Capturing public voices: The role of social media in heritage management. Habitat Int 142:102934. https://doi.org/10.1016/j.habitatint.2023.102934

[5] Liu B (2015) Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press, Cambridge

[6] Nandwani P, Verma R (2021) A review on sentiment analysis and emotion detection from text. Soc Netw Anal Min 11:1–19. https://doi.org/10.1007/s13278-021-00776-6

[7] D'Orazio M, Di Giuseppe E, Bernardini G (2022) Automatic detection of maintenance requests: Comparison of Human Manual Annotation and Sentiment Analysis techniques. Autom Constr 134:104068. https://doi.org/10.1016/j.autcon.2021.104068

[8] D'Orazio M, Di Giuseppe E, Bernardini G (2023) End-user's perception of the state of the building systems through sentiment and emotion analysis methods. In: Fatiguso F, Fiorito F, De Fino M, Cantatore E (eds) Transizione sfide e opportunità per l'ambiente costruito / Transition challenges and opportunities for the built heritage. [EdicoEdizioni, Monfalcone, pp 859–870

[9] Rosin U, Bonazzi M, Casarin F (2022) eWom and sentiment analysis to support decision processes within maritime heritage museum networks. The case of Arca Adriatica. In: Proceedings of 16th International AIMAC Conference on Arts and Cultural Management. Mexico City, pp 1-19

[10] Hutto C, Gilbert E (2014) VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proc Int AAAI Conf Web Soc Media 8:216–225. https://doi.org/10.1609/icwsm.v8i1.14550

[11] Mendes M, Torres-González M, Valença J, Silva A (2022) The Maintenance of Monuments as the Main Trigger to Negative Feelings in Tourists. Buildings 12(12):2153. https://doi.org/10.3390/buildings12122153

[12] Ginzarly M, Srour FJ, Roders AP (2022) The Interplay of Context, Experience, and Emotion at World Heritage Sites: a Qualitative and Machine Learning Approach. Tour Cult Commun 22:321–340. https://doi.org/10.3727/109830421X16345418234065

[13] Valdivia A, Martínez-Cámara E, Chaturvedi I, Luzón MV, Cambria E, Ong Y-S, Herrera F (2020) What do people think about this monument? Understanding negative reviews via deep learning, clustering and descriptive rules. J Ambient Intell Human Comput 11(1):39–52. https://doi.org/10.1007/s12652-018-1150-3
[14] Devlin J, Chang M-W, Lee K, Toutanova K (2019) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Burstein J, Doran C, Solorio T (eds) Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Association for Computational Linguistics, Minneapolis, Minnesota, pp 4171–4186
[15] Loyola M (2018) Big data in building design: A review. J Inf Technol Constr 23:259–284
[16] Bianchi F, Nozza D, Hovy D (2021) FEEL-IT: Emotion and Sentiment Classification for the Italian Language. In: De Clercq O, Balahur A, Sedoc J, Barriere V, Tafreshi S, Buechel S, Hoste V (eds) Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Association for Computational Linguistics, Online, pp 76–83
[17] spaCy (2024) spaCy: Industrial-strength Natural Language Processing. https://spacy.io/. Accessed 8 Jul 2024
[18] nlptown (2020) bert-base-multilingual-uncased-sentiment. Hugging Face. https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment. Accessed 6 Jul 2024
[19] Ekman P (1992) Basic Emotions. In: Encyclopedia of Personality and Individual Differences. pp 1–6
[20] Almeida AMG, Cerri R, Paraiso EC, Mantovani RG, Junior SB (2018) Applying multi-label techniques in emotion identification of short texts. Neurocomputing 320:35–46. https://doi.org/10.1016/j.neucom.2018.08.053
[21] Regione Marche (2024) Patrimonio culturale. https://www.regione.marche.it/Regione-Utile/Cultura/Patrimonio-culturale. Accessed 16 Apr 2024
[22] ICCD (2024) Catalogo generale dei Beni Culturali. https://www.catalogo.beniculturali.it/. Accessed 16 Apr 2024
[23] Mancinelli ML (2018) Gli standard catalografici dell'Istituto Centrale per il Catalogo e la Documentazione. In: Tucci R (ed) Le voci, le opere e le cose. La catalogazione dei beni culturali demoetnoantropologici. [Ministero dei beni e delle attività culturali e del turismo
[24] Istituto centrale per il catalogo e la documentazione, Roma], pp 279–302
[25] dati.cultura.gov.it (2024) dati.cultura.gov.it. https://dati.cultura.gov.it/. Accessed 29 Apr 2024
[26] Picone R (2012) Il rudere architettonico nella storia del restauro. CONFRONTI 0:27–40