www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

LEVERAGING USER-GENERATED CONTENT TO ENHANCE HERITAGE MONITORING PROTOCOLS AND RESOURCES ALLOCATION THROUGH ASPECT-BASED SENTIMENT ANALYSIS

SUBMITTED: October 2024
REVISED: March 2025
PUBLISHED: March 2025
EDITOR: Frédéric Bosché
DOI: 10.36680/j.itcon.2025.017

Marco D'Orazio, Prof.
DICEA, Università Politecnica delle Marche, via Brecce Bianche, Ancona, Italy
ORCID: https://orcid.org/0000-0003-3779-4361
m.dorazio@staff.univpm.it

Elisa di Giuseppe, Prof.
DICEA, Università Politecnica delle Marche, via Brecce Bianche, Ancona, Italy
ORCID: https://orcid.org/0000-0003-2073-1030
e.digiuseppe@staff.univpm.it

Chiara Mariotti, Researcher in Architectural Conservation
DICEA, Università Politecnica delle Marche, via Brecce Bianche, Ancona, Italy
ORCID: https://orcid.org/0000-0002-5496-517X
chiara.mariotti@staff.univpm.it

Maria Francesca Muccioli, PhD student
DICEA, Università Politecnica delle Marche, via Brecce Bianche, Ancona, Italy
ORCID: https://orcid.org/0009-0009-2602-0126
m.muccioli@pm.univpm.it

SUMMARY*: Effective resource allocation for architectural heritage conservation requires reliable and up-to-date data to support proper decision-making. This paper proposes a novel methodology for the integration of reviews continuously generated by users with expert evaluations aimed at supporting decision-makers, which are based on periodic but infrequent site inspections and often fail to capture ongoing issues. This study utilises Natural Language Processing (NLP) and Aspect-Based Sentiment Analysis (ABSA) to improve the monitoring and management of architectural heritage. By leveraging user-generated content from Google Maps, this approach allows the integration of dynamic, real-time feedback into conservation strategies, addressing current limitations. The methodology was applied to data concerning a sample of 70 heritage buildings in Italy's Marche region, whose state of conservation has been catalogued by the Central Institute for Cataloguing and Documentation (ICCD). The analysis highlighted significant discrepancies between expert evaluation and user feedback. While official assessments often miss recent changes or emerging issues, the ABSA-based tool succeeded in capturing insights such as recent restoration interventions or unnoticed maintenance needs. This approach introduces a methodology for a more inclusive, data-driven decision-making process, allowing resource allocation to be more responsive to ongoing conditions. By integrating expert evaluations and public inputs, the proposed methodology allows the adoption of more comprehensive perspectives, detecting blind spots in traditional monitoring systems and enabling more efficient management of resources.*

KEYWORDS*: resources allocation, heritage monitoring, architectural heritage, conservation, user-generated content, user perception.*

REFERENCE*: Marco D'Orazio, Elisa di Giuseppe, Chiara Mariotti & Maria Francesca Muccioli (2025). Leveraging user-generated content to enhance heritage monitoring protocols and resources allocation through aspect-based sentiment analysis. Journal of Information Technology in Construction (ITcon), Vol. 30, pg. 397-417, DOI: 10.36680/j.itcon.2025.017*

1. INTRODUCTION

The efficient allocation of resources in the conservation of architectural heritage depends on the availability of reliable, up-to-date data and information to guide decision-making. Monitoring plays a crucial role in this process, providing information to enable responsible stakeholders to plan or adjust intervention strategies based on emerging needs (Stovel, 2004). Preventive and planned conservation strategies offer, in fact, a more viable alternative to delayed restoration efforts, which often fail to address underlying issues and require substantial public funding (Vandesande and Van Balen, 2018). Ensuring the continuity of such operations is thus essential for the safeguard of built heritage and for the proper management of financial resources (ICOMOS, 2011).

Current protocols rely on periodic site inspections or continuous data recording through sensors (Kioussi *et al.*, 2013), typically conducted by skilled experts within public administrations (D'orazio *et al.*, 2023). While these evaluations are fundamental for the assessment of the state of conservation of assets, they often fail to account for real-time developments, potentially leading to outdated evaluations and suboptimal resource allocation (Masciotta *et al.*, 2021; MiC, n.d.; Rossi and Bournas, 2023).

Despite the recognised importance of participatory approaches, public perceptions are seldom considered in official monitoring processes. Involving the public, however, offers a complementary perspective that can highlight overlooked issues or suggest alternative priorities, fostering a more democratic and inclusive approach to heritage management (Bai *et al.*, 2023; Council of Europe, 2005; Grcheva and Oktay Vehbi, 2021; ICOMOS, 1987, 2008, 2011; Li *et al.*, 2020c, 2020a, 2020b; Miller and Murray, 2018; Rosetti *et al.*, 2022; Sanoff, 1999; UNESCO, 2011; World Heritage Committee, 2015).

Social media platforms now serve as a dynamic repository of user-generated content, offering real-time, unbiased, unstructured data on public perceptions of heritage sites (Foroughi *et al.*, 2023; Jacoby *et al.*, 2024; Torres-Gonzáles *et al.*, 2023). Leveraging this data can complement static reports, reduce "blind spots" and make resource allocation more responsive to current conditions (Amato *et al.*, 2017; Bai *et al.*, 2021, 2023; Monteiro *et al.*, 2015).

Bonci et al. (Bonci *et al.*, 2018) underscore the potential of Artificial Intelligence (AI) tools to enhance decision-making processes and expert assessment, encouraging the use of networked expertise and advanced technologies to support bottom-up dynamics. With the advent of Natural Language Processing (NLP) techniques, analysing large amounts of unstructured data stored on the internet, such as user-generated content, has now become possible (Chowdhury *et al.*, 2021; D'Orazio *et al.*, 2022; Foroughi *et al.*, 2023; Liu, 2015; Quagliarini *et al.*, 2019).

Recently, NLP has gained fortune in the field of building management and maintenance. For instance, D'orazio et al. (D'Orazio *et al.*, 2024) applied NLP algorithms to process 12,000 maintenance requests from end-users for a stock of 23 public buildings, using the data to prioritise interventions and demonstrating the effectiveness of these tools in resource management. Similarly, Bugalia er al. (Bugalia *et al.*, 2022) applied NLP and Machine Learning (ML) to classify worker-reported safety notices from a construction site in Kuwait, highlighting the potential of these tools for improving decision-making based on unstructured data.

Among NLP applications, as suggested by Loyola (Loyola, 2018), one of the most promising technologies for managing big data is Sentiment Analysis (SA) (Fang and Zhan, 2015; Liu, 2015).

SA, or opinion mining, is the computational study of people's opinions, sentiments and emotions. SA is fundamental for leveraging user-generated data online, allowing for sentiment classification (positive, negative, or neutral) and more detailed analyses, such as assessing sentiment intensity or, in the case of emotion analysis, identifying specific emotions (e.g., joy, anger, fear, or sadness, depending on the model used) (Liu, 2015; Medhat *et al.*, 2014).

SA is commonly applied to measure user satisfaction with products and services; an example is given by Chowdhury et al. (Chowdhury *et al.*, 2021) who applied SA to 1152 user reviews of web-based collaborative tools developed for the construction sector to include in their performance evaluation users' feedback.

Although SA use is expanding in other fields such as politics, healthcare, and, more recently, building maintenance and architectural heritage preservation (Liu, 2015; Nazir *et al.*, 2022; Wankhade *et al.*, 2022).

D'Orazio et al. (D'Orazio *et al.*, 2022) conducted a comparative analysis of the efficacy of various SA techniques in processing requests from computerised maintenance management systems for university-administered buildings, thereby demonstrating the potential of the tool in prioritising interventions. Valença et al. (Valença *et*

al., 2024) used SA to survey inhabitants of Bairro de Alvalade (Lisbon) regarding façade conditions and found a strong correlation between public perception and objective indicators of conservation, such as decay detection via multispectral cameras.

Several studies have furthermore demonstrated the applicability of SA to user-generated content for enhancing built heritage management.

In their study, Amato et al. (Amato *et al.*, 2017) employed SA as a ranking criterion for the Smart Context-aware Browsing Assistant for Cultural Environments (SCRABS) tool. This tool is designed to assist users in navigating heritage information across a range of heterogeneous data sources, including digital libraries, social networks, and other web services. Mendes et al. (Mendes *et al.*, 2022) used VADER (Valence Aware Dictionary for Sentiment Reasoning) (Hutto and Gilbert, 2014) to analyse TripAdvisor reviews of two Iberian Peninsula monuments, identifying maintenance actions (e.g., scaffolding) as the primary triggers of negative sentiment. Rosin et al. (Rosin *et al.*, 2022) applied VADER to reviews from TripAdvisor and Google Business for the Arca Adriatica maritime museums network to assess how electronic word-of-mouth (eWOM) could improve management, user experience, and marketing strategies.

Valdivia et al. (Valdivia *et al.*, 2020) conducted Aspect-Based Sentiment Analysis (ABSA) on TripAdvisor reviews of three Spanish monuments to detect opinion polarity and focused on results that showed a negative sentiment. Further analysis of individual reviews identified the aspects requiring action from cultural managers.

As user-generated content often contains opinions on multiple aspects within a single sentence, it is crucial to detect the target of analysis to ensure reliable results.

SA can be developed at different levels, specifically at document level to extract the general sentiment expressed by a whole document, sentence level to discover whether a single phrase expresses a positive, negative, or neutral sentiment, or aspect level. ABSA focuses on identifying exactly what users appreciate or criticise, detecting opinion targets within sentences to provide the depth of analysis required for practical applications (Liu, 2015; Nazir *et al.*, 2022; Zhang *et al.*, 2023). For instance, in the sentence *"Apple is doing well in this poor economy",* positive sentiment would be assigned to "Apple" while negative sentiment would be associated with "economy", returning a nuanced sentiment "summary" for the sentence (Liu, 2015).

ABSA remains a challenging field due to complex linguistic patterns, difficulties in interpreting factual sentences that imply opinions, and the lack of sufficient training data for specific domains. Existing systems struggle with complex sentences, negation, modality, and accurate opinion target identification. Moreover, each domain requires unique lexicons and methods, complicating cross-domain classification accuracy (Liu, 2015; Nazir *et al.*, 2022).

The introduction of transformer-based pre-trained NLP models such as BERT (Devlin *et al.*, 2019) and RoBERTa (Liu *et al.*, 2019) has marked a significant advancement in this field. These models improved specific-design deep learning models (Zhang *et al.*, 2023) and revolutionised the world of NLP (Jain, 2022; Pourkeyvan *et al.*, 2024; Ravichandiran, 2021).

BERT, which stands for Bidirectional Encoder Representations from Transformers, was introduced by Google in 2017 (Devlin *et al.*, 2019). Like other Transformer models (Vaswani *et al.*, 2017), it features an encoder based on self-attention mechanisms, allowing it to better capture word relationships and contextual information within sentences. Its innovations, including Masked Language Modelling (MLM) and Next Sentence Prediction (NSP), have greatly enhanced the model's ability to understand context. Once pre-trained, BERT can be fine-tuned for specific tasks like translation, summarisation, or sentiment classification (Jain, 2022; Pourkeyvan *et al.*, 2024; Ravichandiran, 2021).

These advancements present new opportunities for improving heritage management and decision-making through the leveraging of unstructured user-generated content.

2. RESEARCH AIM

The present study aims to leverage pre-trained BERT-based NLP models to perform ABSA on user-generated content from social network platforms concerning architectural heritage selected from public administration records. The primary goal is to develop a standardised methodology for extracting user perceptions specifically related to heritage, focusing on architectural aspects such as the physical consistency of the site, its surrounding

environment (e.g. gardens), and the possible presence of valuable objects (e.g., furniture, artworks) inside, as these may contribute to the overall perception of the asset and its value. These factors were isolated from unrelated elements like the building's function, ensuring the analysis captures the architectural experience more accurately. This addresses the current lack of protocols for analysing targeted, topic-specific user feedback.

The study has twofold objectives, reflected in the following hypotheses*:*

- *Hypothesis 1*: ABSA can effectively isolate heritage-specific aspects from user-generated content, minimising the influence of non-pertinent factors. This will be demonstrated by comparing the distribution of user-assigned ratings on social media, which reflect a global evaluation across multiple aspects, with ABSA-derived scores. While the overall perception may show a similar trend, slight differences are expected due to ABSA's focus on architectural aspects.
- *Hypothesis 2*: Insights provided by the ABSA methodology can enhance existing heritage monitoring protocols, supporting experts in detecting criticalities and improving decision-making and resource allocation. Discrepancies between expert evaluations of heritage conservation states and ABSA results may reveal overlooked issues or areas where public perception offers unique, actionable insights.

The case study focuses on data concerning historic buildings in the Marche region of Italy, with monuments selected from the Central Institute for Cataloging and Documentation (ICCD, "Istituto Centrale per il Catalogo e la Documentazione") records. User-generated content, such as reviews from Google Maps, was collected for analysis.

The ABSA methodology, further detailed in the following sections, involves extracting relevant heritage-related words from reviews, segmenting the texts, applying a BERT-based algorithm to analyse sentiment, and calculating an overall sentiment score for each monument.

The paper is organised as follows: Sect. 3 introduces the case study and data sources; Sect. 4 details the proposed methodology; Sect. 5 presents the ABSA results and evaluates the hypotheses through comparisons with user-assigned ratings and official assessments; Sect. 6 provides insights into the potential implications of this research for future studies and real-world applications.

3. CASE STUDY

The selected case study on which to test the proposed methodology is based on data concerning heritage architectures located in the Italian region of Marche.

Italy is home to a wealth of historic cities and buildings, boasting the largest number of UNESCO heritage sites worldwide (UNESCO, n.d.; World Economic Forum, 2020). The region Marche itself is rich in architectural heritage, counting more than 1000 monuments, 106 castles, 15 fortresses, thousands of churches and 72 theatres. It also has the highest museums-to-population ratio in Italy (Lorenzini *et al.*, 2011; Marche, n.d.). The region includes one UNESCO World Heritage site, the historic centre of Urbino (MiC, n.d.), and two UNESCO Creative Cities (UNESCO, n.d.): Pesaro for music and Fabriano for crafts and folk art (Letsmarche, n.d.).

In Italy, the Ministry of Culture (MiC) is the authority responsible for protecting cultural heritage (MiC, n.d.), operating through specific peripheral offices (MiC, n.d.; Presidenza del Consiglio dei Ministri, 2021). Among these, the Superintendencies for Archaeology, Fine Arts, and Landscape (SABAP, "Soprintendenze Archeologia, Belle Arti e Paesaggio") oversee the conservation and promotion of heritage sites. In the Marche region, the SABAP-AN-PU serves the provinces of Ancona (AN) and Pesaro and Urbino (PU), while the SABAP-AP-FM-MC covers the provinces of Ascoli Piceno (AP), Fermo (FM), and Macerata (MC). These official bodies manage heritage conservation on the basis of comprehensive knowledge of the cultural assets in their respective areas of competence.

To manage the extensive number of heritage sites, several platforms have been developed to collect information on protected architectures, such as "Vincoli in Rete" (Constraints on the Net) (MiC, DGABAP, ICR, n.d.), "Carta del Rischio" (Risk Chart) (MiC; ICR, n.d.), and the "Catalogo Generale dei Beni Culturali" (General Catalogue of Cultural Heritage) (ICCD, n.d.). The latter, which stores data in the SIGECweb system (ICCD, n.d.), is particularly relevant as it organises information collected by the Central Institute for Cataloguing and Documentation (ICCD) (ICCD, n.d.).

ICCD is the authority responsible for the collection of data derived from cataloguing campaigns conducted by local authorities (superintendence, provinces, etc.) (ICCD, 2023, n.d.), as well as their organisation and control. These campaigns document the characteristics of heritage objects, including their state of conservation at the time of inspection, a feature that sets the platform apart from others in Italy (Mancinelli, 2018).

The information is compiled into structured catalogue sheets based on standardised cataloguing systems (ICCD, n.d.), which cover not only built heritage but also works of art, archaeological artefacts, musical instruments, historic clothing and so on (Mancinelli, 2018). For this study, only "A" sheets (those specific to architecture) were considered (ICCD, 2015; Mancinelli, 2018). These catalogued sheets for the Marche region form the initial data source for the research, including both administrative data and the architectural objects analysed in the ABSA. The complete dataset for Marche's heritage architecture was downloaded in .json format from dati.beniculturali.it (ICCD, n.d.).

The data were then cleaned to retain only essential information for the study. Irrelevant details, such as land registry data or bibliography, were excluded. The final dataset contained the following information for each building: unique national identification code (ICCD, 2015; Mancinelli, 2018), denomination, location (latitude and longitude), state of conservation, additional notes on the state of conservation (e.g. causes of decay), and the dates of both compilation and update of the record. A total of 6835 entries on architectural data were retrieved and used as the foundation for the subsequent analysis.

4. MATERIALS AND METHODS

The ABSA procedure can be summarised in three main steps:

- Extraction and organisation of user-generated content (Sect. 4.1): user reviews related to the architectural sites identified in the case study (from the "A" cataloguing sheets) were collected from Google Maps.
- NLP-based analysis (Sect. 4.2): the NLP tool spaCy was used to detect relevant heritage-related words (tokens) within the user-generated texts. These tokens served as "targets" for the ABSA, allowing for the identification of relevant sentence segments for further analysis.
- Sentiment Analysis (Sect. 4.3): a BERT-based model was applied to perform SA on the identified text segments, which were then aggregated to generate a sentiment score for each heritage object.

These steps will be detailed in the following paragraphs and are illustrated in Figure 1.

Figure 1: Representation of the workflow for the proposed methodology.

4.1 Extraction of user-generated content

In the initial phase of the study, architectural sites listed in the ICCD catalogue were selected based on their designation as public property, thus ensuring easier access for users. Most of the heritage recorded in the catalogue is privately owned, restricting visitor access and, consequently, user-generated content.

Google Maps was chosen as the platform for data extraction. It offers a vast repository of user-submitted information, including 5-star ratings, optional written reviews, and photos documenting visitor experiences. As one of the most popular review platforms, Google Maps has seen significant growth in recent years and has become a primary source for real-time information on various locations, including cultural sites (Chen and Chang, 2024; Mathayomchan and Taecharungroj, 2020; Murphy, 2018).

Unlike other platforms, Google reviews are not verified through content control procedures, allowing for the inclusion of unfiltered opinions. While this can lead to issues with spam or fake reviews (Liu, 2015) –which can be reported and removed if they violate community guidelines (Google, n.d.)– it provides a larger and more diverse data set, allowing for a more robust testing of the methodology.

To match the user-generated content with the catalogued heritage sites, manual searches were conducted on Google Maps. The application of automated matching methods proved to be ineffective due to a number of issues, including the presence of inconsistent naming conventions (e.g. the use of abbreviations such as "S." for "Saint") and the occurrence of name changes over time (e.g. the former "bishop's seminary" in Urbino being transformed into the "Albergo S. Domenico").

Geographic matching by coordinates was also problematic due to errors in the manual entry of latitude and longitude in the catalogue. Setting a tolerance radius around mapped points was impractical in historic city centres, where buildings are closely spaced, increasing the likelihood of mismatches (see Fig. 2).

Once the relevant heritage sites were identified, those with user-generated reviews were selected for data extraction. Using the Google Chrome extension Instant Data Scraper (UAB Webrobots, n.d., n.d.), information from each review was collected, including username, posting date, rating, and review text, which were stored in a .json file. To protect user privacy and copyrights, usernames were anonymised and review texts are not presented in the original language.

A total of 70 heritage sites were included in the study, with each having its own .json file. These files were named using the unique code from the ICCD's catalogue sheets (Mancinelli, 2018), allowing for precise matching between the two datasets. The complete list of items is provided in the supplementary material (Supp. 1).

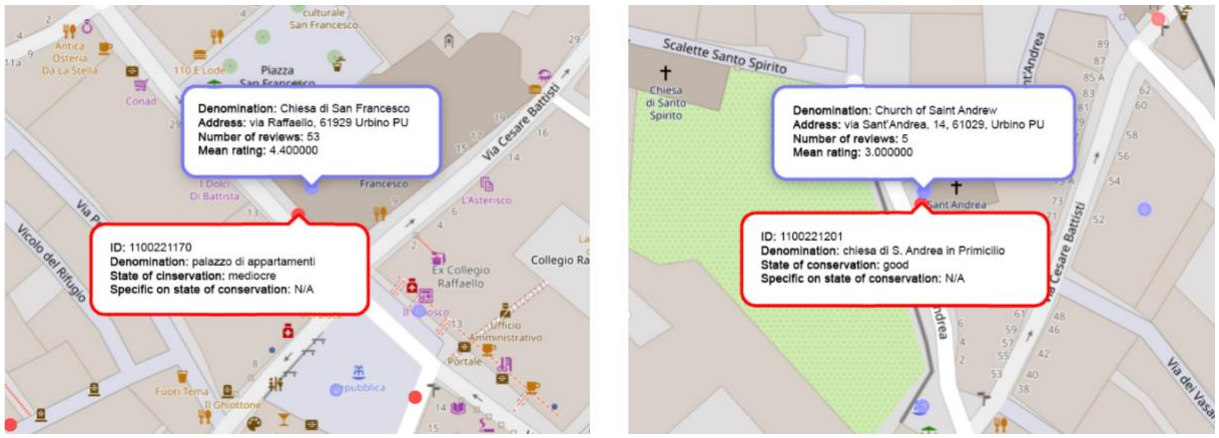

Figure 2: Examples of cases of mismatch (red dots refer to ICCD-mapped items, violet dots to Google Maps-retrieved ones): on the left, the potential error in pairing the objects is due to buildings' proximity in historic city centres, not allowing a match by coordinates; on the right, the impossibility of performing the match is due to the different denomination.

The significant reduction in the number of identified heritage items on Google Maps, from the initial 6835 catalogued items to 70, can be attributed to two factors. First, the majority of catalogued heritage is privately owned, limiting the number of accessible sites. Second, many of the catalogued items are considered "minor" heritage, which may not stimulate the public's interest in visiting the site and leaving a review. This discrepancy highlights that the ICCD catalogue includes a broad definition of "heritage", not always recognised by the general public, particularly those without expertise in heritage preservation. Additionally, in accordance with Italian legislation, all publicly owned architectural assets that have been in existence for a minimum of 70 years are

automatically included on the list of protected properties (*D.Lgs. 42/2004 (Bb.Cc.Aa.)*, 2004) and consequently registered in the ICCD catalogue. This regulation increases the likelihood of including more recent heritage (such as schools or public offices built around the 1950s), whose value is not always recognised by the general public.

The use of Google-translated text for SA is a well-established practice in the literature (Foroughi *et al.*, 2023; Li *et al.*, 2022; Sergiacomi *et al.*, 2022): in this study, the majority of reviews were written in Italian, while a few in other languages were automatically translated by Google to ensure consistency in the dataset. Although multilingual SA is possible (Abbasi *et al.*, 2008; Dashtipour *et al.*, 2016; Mountstephens and Quen, 2023), this study focused solely on reviews in Italian, either in their original form or translated, to maintain uniformity.

4.2 Individuation of heritage related words and segmentation of reviews' texts

This phase of the procedure consists of two sub-steps: identifying heritage-related words in the previously extracted review texts and isolating sentence segments containing these words. For this task, the Python library spaCy ("Spacy.Io", n.d.) was employed.

SpaCy is an open-source NLP library for Python that supports more than 75 languages. The Italian language is manageable through the use of specific libraries, such as it_core_news_sm (Explosion, 2024). Among its features, it can break sentences into simple units (tokens) and use Part of Speech (POS) tagging to assign labels to each token, such as noun, adjective, verb, etc. This facilitates an analysis of the syntactic structure and an understanding of the roles that words play in a sentence.

By combining spaCy with Python's default Collection module, words functioning as subjects or direct objects within the sentences were identified and counted, then ranked from most to least frequent. These words represent elements directly associated with expressions of appreciation or dissatisfaction.

Next, the identified lexical items were manually filtered to retain only those relevant to the heritage domain; an example is provided in Table 1.

Table 1: Example of words (subject and direct objects) retrieved by SpaCy for the church of Portone in Senigallia and example of the selection of terms related to heritage. This example was selected for its efficacy in illustrating the process of token selection: words related to the building (church, chapel, architecture) were kept, as well as terms related to the secondary heritage that may be contained in it (tabernacle, works); other terms such as "pena" (sorrow) were kept because of the Italian expressions such as "fare pena" (to suck) or "vale la pena" (it is worth) that may appear in sentences. Terms relating to the religious sphere, such as "priest", "punishment" or "peace", were clearly excluded since they do not concern our interest.

Heritage building	**ICCD code**	**List of word (subject and direct object) with frequency (within brackets) – English translation by authors**
Parrocchia Chiesa del Portone	1100219443	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)
		Selected words
		Building related: church; chapel; architecture Works of art/furniture: tabernacle; works Users' expressions: atmosphere; sorrow; favourite; enchanting

Although it was not possible to establish a universal rule for token selection due to the diverse nature of the architecture in question, the following criteria were applied: (i) words related to architectural structures or their surroundings (e.g. gardens), (ii) terms referring to historically significant furniture or artworks inside the buildings (as they contribute to the overall perception of the site's value) and (iii) expressions related to visitor actions or experiences (e.g. the term "visit," which helps capture sentences like "worth a visit").

Once the tokens were identified and listed for each architecture, they were used to select the relevant segments of review texts for SA. Segments were defined as portions of text separated by punctuation marks or conjunctions, as these often signal the boundaries of individual thoughts or ideas.

Using spaCy on a dataset containing all textual reviews for the heritage sites, punctuation marks and conjunctions were identified as separators. This allowed the segmentation of reviews into distinct parts, enabling the differentiation between sections related to heritage and other unrelated content (see Table 2).

Table 2: Example of review texts for the church of Portone in Senigallia.

Heritage building	ICCD code	Selected tokens for the individuation of segments – English translation by authors	User	Example of review – English translation by authors
Parrocchia Chiesa del Portone	1100219443	church; atmosphere; sorrow; chapel; tabernacle; architecture; works; 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"

4.3 Sentiment analysis

Once the heritage-related segments within each review were identified, enabling the application of an aspect-based approach, they were subjected to SA using the pre-trained model bert-base-multilingual-uncased-sentiment (nlptown, 2020), retrieved from the Hugging Face platform, which offers a wide range of pre-trained models (Hugging Face, n.d.). As the name suggests, the model is based on BERT, supports multiple languages (including Italian), and does not distinguish between uppercase and lowercase letters. The model was chosen for its ability to provide sentiment scores on a 5-star scale, consistent with the ratings given by users, alongside a confidence percentage representing the algorithm's certainty in assigning the score. For the Italian language, using user-assigned star ratings as the ground truth, the model demonstrated an accuracy of 59%, which rose to 95% when the model's predictions differed by no more than one star from the user's rating (nlptown, 2020).

Each review segment was analysed, and a 5-star score was assigned to each. To calculate a single sentiment score for the entire review, the average of the segment scores was computed. Once all reviews for a heritage site were analysed and individual scores assigned, the overall sentiment score for the monument was determined by averaging the scores of all reviews (see Fig. 3).

Figure 3: Schematic representation of the process for obtaining a single sentiment score for a heritage architecture.

Scores of 1 and 2 were classified as negative, 3 as neutral, and 4 and 5 as positive, as the focus was on the overall perception of heritage.

5. RESULTS AND DISCUSSION

Following the calculation of ABSA scores for all 70 heritage sites included in the case study (presented in Section 4.1), the results were analysed to assess the validity of the hypotheses defined in Section 2.

First, the distribution of user-assigned ratings on Google Maps and ABSA-derived scores is compared to evaluate hypothesis 1. This comparison aims to highligh differences attributable to ABSA's ability to isolate heritage-specific aspects. While overall trends may align, slight discrepancies are expected due to the aspect-based approach (Section 5.1).

Second, ABSA results are compared with expert evaluations of the conservation states of the heritage sites to test hypothesis 2. Potential discrepancies between official assessments and user perceptions may highlight overlooked issues or criticalities, demonstrating ABSA's value in enhancing monitoring protocols (Section 5.2).

5.1 Hypothesis 1: analysing differences between user-assigned ratings and ABSA scores

The distribution of ratings directly provided by users (user-assigned rating) and the results from the ABSA (ABSA scores) were analysed and compared using histograms (Fig. 4). Both graphs show a similar distribution pattern, demonstrating the effectiveness of the selected SA model in capturing the overall perception of users from their written reviews. A generally positive perception of built heritage emerged, with the majority of ratings at 4 stars.

However, ABSA results showed fewer 5- and 1-star ratings compared to user-assigned ratings, while 2- and 3-star ratings were higher. This discrepancy may be attributed to the aspect-based approach focused specifically on heritage, as users may have based their ratings on multiple factors beyond the architectural aspect alone.

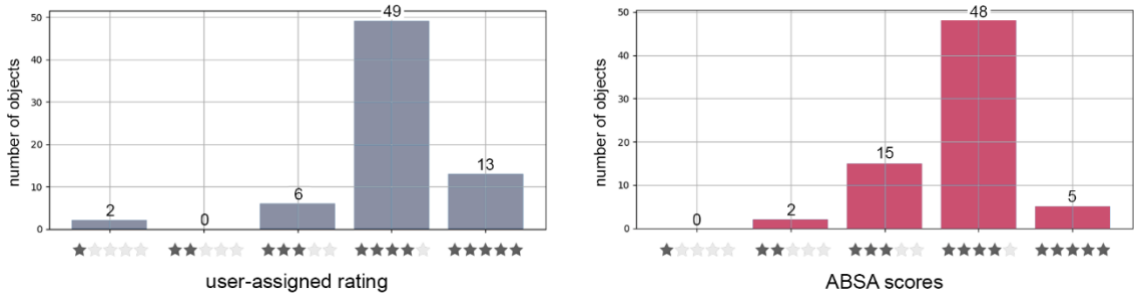

Figure 4: Histograms representing the distribution of user-assigned ratings (left) and the distribution of the resulting ABSA scores (right).

Alongside the general evaluation of the entire dataset, a more in-depth analysis categorised the selected heritage objects based on their use: directional (4 buildings), fortifications (5), education (4), monuments (1), museums (4), productive (1), religious (39), receptive (8), services (1), theatres (2), and private villas (1). The first notable

observation from this breakdown is the predominance of religious buildings, particularly churches and monasteries, which reflects both their abundance in the region and the strong visitor interest in this type of architecture.

Once again, user-assigned ratings (only those linked to textual reviews, with non-textual ratings excluded from the mean) were compared with ABSA results to identify discrepancies (corresponding histograms can be found in supplementary material Supp. 2). The most significant discrepancies, where no alignment was observed between user ratings and ABSA scores, occurred in the categories of productive, services, and private villas—each represented by only one item (Fig. 5). These items are the former Mancini Furnace ("ex fornace Mancini") in Pesaro, the Mail Palace ("palazzo delle Poste") in Ancona, and Villa Vismara in Pesaro (Fig. 6).

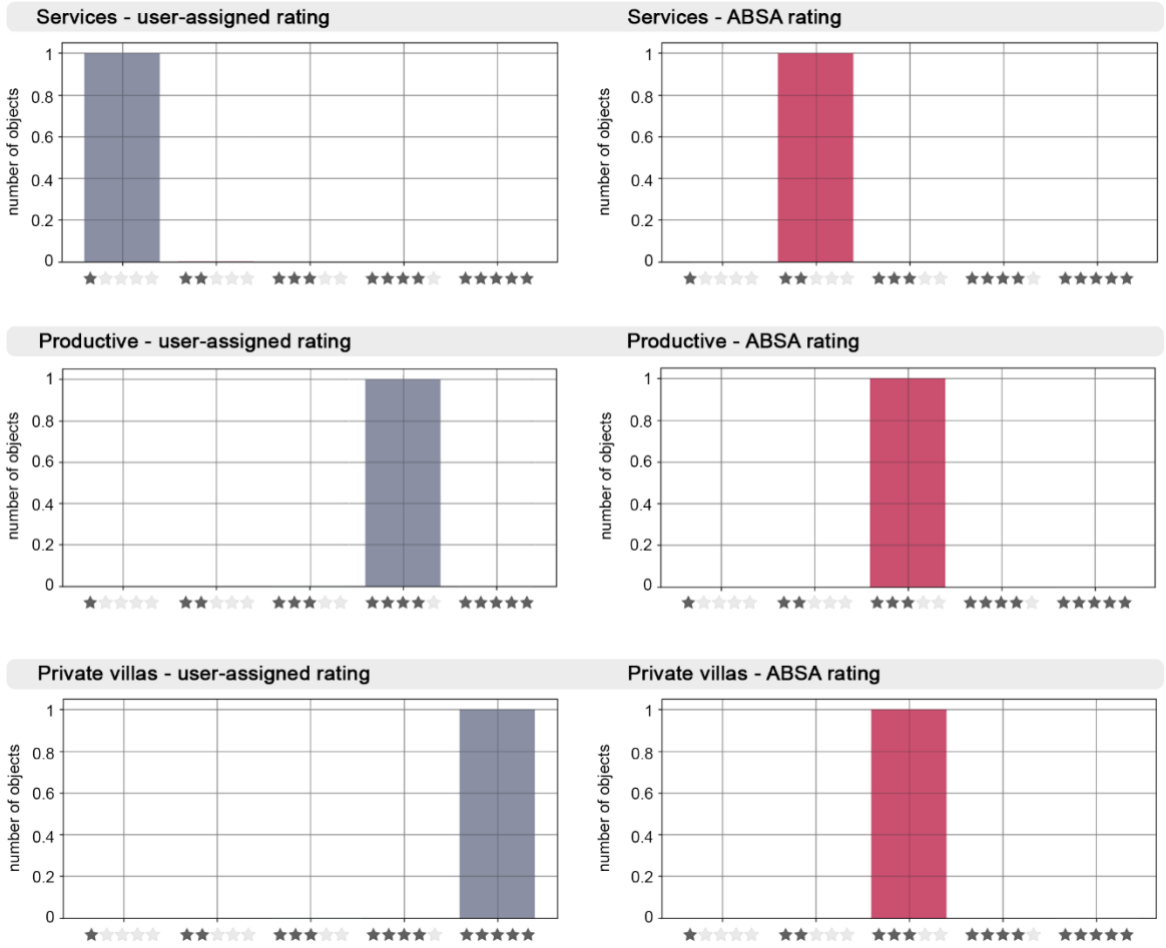

Figure 5: Histograms representing the distribution of ratings directly assigned by users (left) and the distribution of the resulting scores of the ABSA (right) for Services, Productive and Private villas categories.

Starting with the Mail Palace, which had the lowest rating, the shift from a 1-star user rating to a 2-star ABSA score can be attributed to the aspect-based approach. However, a closer examination of the text reveals that, even when selecting target words (e.g., "war" and "place"), the reviews primarily address complaints about the postal service, which is not relevant to this study.

In contrast, the discrepancies for the Mancini Furnace and Villa Vismara are more complex. Both buildings received high user ratings, averaging respectively 4 and 5 stars, while the ABSA yielded more neutral evaluations of 3 stars. For the Furnace, there were two reviews: one user gave it a 3-star rating, citing the need for redevelopment, while the other, alongside a 5-star rating, described the experience of entering an abandoned site. The latter review highlights the appeal of ruins and abandonment, which likely explains the divergence between the user's high rating and the ABSA's neutral interpretation of the text.

For Villa Vismara, three reviews were given, all with 5-star ratings. Upon analysing the textual content, two reviews clearly express appreciation, while the third simply asks about the building's opening status (opening issues also appear in another review). The SA model interpreted this question as a neutral, objective statement, which led to a lower overall score. The corresponding reviews are presented in Table 3.

Table 3: Reviews for ex fornace Mancini and Villa Vismara.

Heritage building	User	User-assigned rating	Text of review – English translation by authors
Ex fornace Mancini	User_01	5/5	Together with hiddenname01 and hiddenname02, I went to explore the disused Mancini furnace. The plants - unfailing brambles - laid siege to 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 pass through it and we are inside.
	User_02	3/5	A ruin awaiting targeted redevelopment.
Villa Vismara	User_01	5/5	Private villa, not open to the public. Very beautiful!
	User_02	5/5	Is it open?
	User_03	5/5	Nice place

Figure 6: Left image: the former Mancini furnace, Pesaro, belonging to Productive category; right image: Vismara Villa, Pesaro, Private villas category.

In addition to using histograms for analysing the distribution of average user ratings and ABSA scores for each heritage item, boxplots were also employed. These graphs illustrate the distribution of ratings and ABSA results for individual reviews, providing insights into the variability and range of evaluations. Both graphs indicate an overall positive perception of heritage items, although the ABSA results show a slight tendency towards more negative and dispersed outcomes. This suggests that the ABSA tool captures a broader range of opinions, potentially less influenced by extreme views compared to user ratings. This observation is supported by the greater presence of outliers in the rating graph. A more in-depth analysis, categorised by the destination of use, further highlights the wider variability of ABSA results, which tend to skew towards lower scores. Both the general and categorised representations are included in the supplementary material attached to this article (Supp. 3).

5.2 Hypothesis 2: discrepancies between user perceptions and expert judgements

Final considerations involved comparing user ratings and ABSA results with the state of conservation recorded in the ICCD catalogue sheets to evaluate the utility of the proposed methodology in supporting professionals in heritage management and detecting criticalities. This comparison allowed for an evaluation of user perceptions alongside expert assessments by heritage technicians.

For the purposes of this analysis, the judgements on the state of conservation, user ratings, and ABSA results were grouped into three categories reflecting sentiment polarity (positive, neutral, and negative), without specifying

intensity. Classifying the expert judgements, however, was more complex due to the subjective nature of the terms used to describe the state of conservation. No standardised guidelines define these terms, making it relatively straightforward to distinguish between "good" and "bad" states of conservation, but more challenging to differentiate terms like "minor damage" and "discrete". The adopted classification is detailed in Table 4: Positive labels were straightforward, including only "good" and "excellent." Descriptions indicating ongoing work were marked as neutral, as were terms like "discrete", which did not explicitly convey a positive or negative sentiment. The negative category encompassed terms referring to damage and collapse, as well as "bad", "very bad" and "mediocre". The term "ruin" was also included in this category, interpreted as a synonym for "collapsed" in the context of a conservation assessment. However, this interpretation is open to debate, as a ruin may have cultural or historical significance and be in an optimal state of conservation despite its dilapidated appearance (Picone, 2012).

Table 4: Categorisation into positive, negative and neutral of the ICCD recorded state of conservation, users-assigned rating and results of ABSA.

Polarisation	**Colour**	**ICCD state of conservation**	**User-assigned rating**	**ABSA score**
Positive	Green	Excellent	4, 5	4, 5
		Good		
Neutral	Yellow	Discrete	3	3
		Absence of damage		
		Work in progress		
		Under renovation		
Negative	Red	Minor damage	1, 2	1, 2
		Moderate damage		
		Medium damage		
		Damage of moderate level		
		Mediocre		
		Bad		
		Very bad		
		Severe damage		
		Very severe damage		
		Partly collapsed		
		Ruin		
		Collapse		
		Collapsed		

The distribution of heritage into negative, neutral, and positive categories (based on technician-assigned judgements of conservation, user ratings, and sentiment scores from the ABSA) was analysed and recorded in Figure 7. A quick examination of the data reveals that the most negative assessments come from the catalogue reports, where the number of negative evaluations is nearly equal to the positive ones.

In contrast, the distributions reflecting user perceptions (user-assigned ratings and ABSA results) show a significantly more positive outlook, with the majority of items falling into the positive category, far outnumbering

the neutral and negative ones. The differences between user ratings and ABSA results, which can be attributed to the more nuanced nature of the SA, have already been discussed.

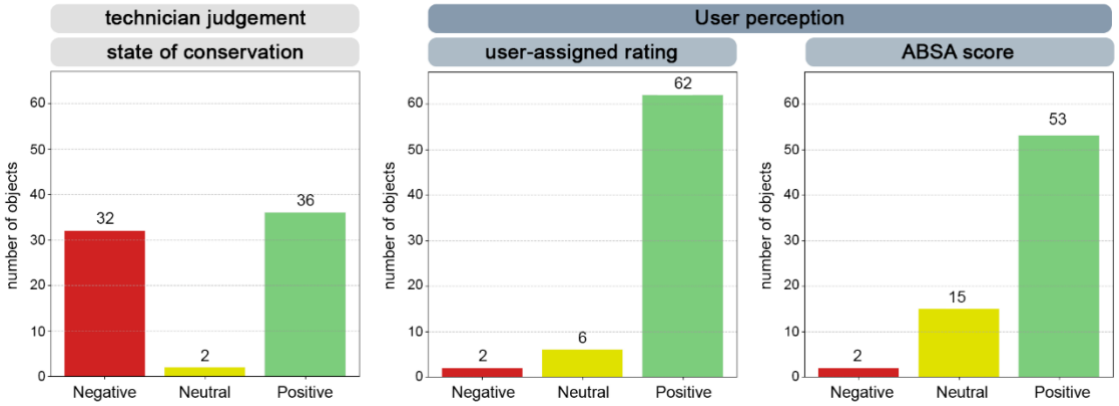

Figure 7: Histograms reporting the distribution of heritage buildings into positive, negative and neutral categories, according to Table 4: technician-assigned judgement on the state of conservation (left), user-assigned ratings (centre) and resulting score of the ABSA (right).

Subsequently, the selected heritage sites were geographically mapped based on the coordinates extracted from the catalogue sheets, with each site assigned a colour corresponding to both the expert evaluations and user perceptions (Fig. 8).

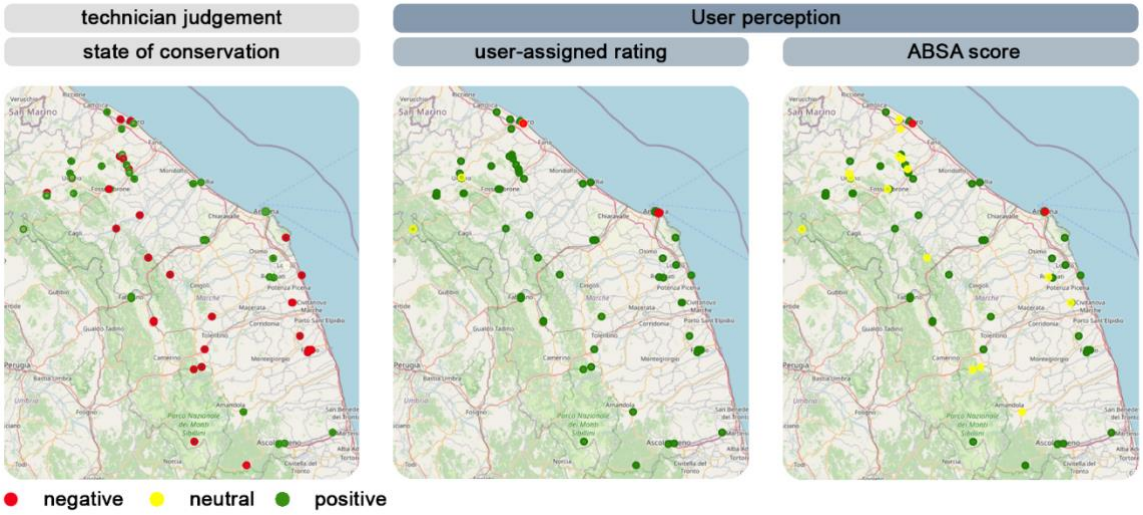

Figure 8: Maps picturing the position of heritage on the regional territory and their corresponding ICCD state of conservation (left), rating directly assigned by users (centre) and resulting scores of ABSA (right). Green markers represent positive connotations, yellow markers neutral and red markers negative ones, according to Table 4.

Technician evaluations depict a more negative conservation state, particularly in the southern areas and hinterlands, while user ratings and ABSA results portray a generally more optimistic view. The ABSA results, however, show greater variability, with more neutral cases scattered across the region, underscoring its ability to capture diverse perceptions.

An analysis of the notes in the "Specific Indications" section (STCO) of the ICCD catalogue sheets, providing additional context for the "State of Conservation" (STCC) assessment, revealed that most of the negative cases were related to damage caused by the 2016 earthquake in central Italy (Reluis, n.d.). A closer examination of individual cases revealed instances where the state of conservation was recorded as negative in the catalogue sheets, but user perceptions were found to be positive due to restoration and refurbishment projects carried out

after the sheets were compiled or updated. This discrepancy highlights the need for more frequent cataloguing campaigns (Fig. 9).

Conversely, two cases emerged where a positive technical assessment was met with negative user ratings and ABSA scores. These cases involved the Mail Palace in Ancona and the city cemetery of Pesaro. In the case of the Mail Palace, the negative perception stemmed from the building function, as discussed in the previous Sect. 5.1. For the Pesaro cemetery, an analysis of individual reviews revealed reports of structural damage and the need for maintenance (Table 5). The corresponding catalogue sheet, dated 2004, further underscores the need for updated assessments.

Figure 9: Castel di Luco (di Luco Castle), in Acquasanta Terme (Ascoli Piceno), ICCD code number 1100060612, date of compiling 1994, date of updating 2018: the castle under works of repair of damages and refurbishment, started after 2018 (left), and the castle today (right).

Table 5: Reviews for the city cemetery of Pesaro, selected among the ones pointing at damages in the complex and highlighting the necessity of maintenance interventions.

Heritage building	User	User-assigned rating	Text of review – English translation by authors
Cimitero San Decenzio	User_01	1/5	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)
	User_02	1/5	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 A disgrace of a poorly constructed place with the new pavilions in a pitiful state, after the charge they demand for burials there needs to be an internal shuttle to reach the various entrances from the many areas that are cordoned off due to the risk of falling materials.
	User_03	1/5	Is it open?
	User_06	5/5	Central, expanding but poorly maintained cemetery. The old part is very much in decay.

6. CONCLUSIONS

The continuous care of architectural heritage, based on a “preventive” rather than “curative” approach (Moioli *et al.*, 2014; Della Torre, 2023), and supported by regular monitoring, is of crucial importance for the proper conservation of cultural heritage and allocation of resources.

This study presents a novel methodology that utilises NLP and ABSA to improve the monitoring and management of architectural heritage. By leveraging user-generated content from Google Maps this approach allows the integration of dynamic, real-time feedback into conservation strategies, addressing current limitations in traditional monitoring protocols that often rely on outdated, static catalogue sheets. By systematically identifying discrepancies between expert and user evaluations, this study demonstrates how integrating user-generated feedback can inform the prioritisation of restoration efforts and the refinement of conservation strategies. This methodology provides a framework for evidence-based decision-making and ensures that resources are allocated more efficiently, addressing real-time needs highlighted by public perceptions.

The case study focused on reviews of 70 heritage architectures in Italy's Marche region, selected from the ones in the General Catalogue of Cultural Heritage, managed by the ICCD. The data were subjected to an ABSA to focus specifically on architectural aspects, excluding comments unrelated to the monument itself.

The first key result of the ABSA methodology showcased an overall positive perception of the architectural heritage in the Marche region. In accordance with the hypothesis 1 formulated in Sect. 2, compared to directly user-assigned 5-star ratings, the model returned more nuanced perceptions and slightly lower scores. This is likely due to the aspect-based approach, as users may have given more "extreme" ratings also based on non-architectural factors, while written reviews tend to reflect a broader range of opinions.

Secondly, an in-depth analysis of specific monuments gave insights on user preferences, such as the appeal of ruined or abandoned architectures.

A novel aspect of this study is the comparison between expert assessments, recorded in ICCD catalogue sheets, and user perceptions from online reviews. The comparison thus highlighted discrepancies: in some cases, expert-assigned negative evaluations were met with positive user feedback, often for sites that had undergone restoration following the 2016 earthquake that were not recorded in the catalogue. Conversely, there were instances where users complained about damages to assets that had been rated positively by experts in catalogue sheets dating back up to 20 years. This analysis underscores the outdated nature of some catalogue records and illustrates how discrepancies between expert assessments and user feedback can lead to misallocated resources. For instance, funding may have been reserved for earthquake-affected monuments that no longer require attention, while critical sites—such as the cemetery noted in this study—remain overlooked despite evident public concern.

This study introduces a new, straightforward methodology for conducting ABSA on unstructured data from social media platforms. The analysed heritage assets were matched to entries in the ICCD catalogue, allowing for a comparison of expert and user perspectives. From a technical standpoint, the potential of the model was confirmed, and critical gaps in current monitoring strategies were identified, particularly the need for more frequent archiving campaigns. Discrepancies identified in the comparison reveal how outdated catalogue records can lead to resource misallocation, with funds directed to sites that no longer require attention while more critical assets remain underfunded. By integrating ABSA with user-generated feedback, this approach bridges the gap between static records and real-time needs, ensuring that restoration priorities better align with evolving public perceptions and site conditions, thereby supporting evidence-based decision-making, in accordance with hypothesis 2 (Sect. 2).

Future research could focus on refining the model, especially in selecting target words and identifying relevant text segments. More efficient NLP models could be developed or adapted to better segment text, incorporating rules grounded in Italian grammar to enhance precision. Additionally, the development of a system to automatically compare catalogue data with user-generated content on platforms like Google Maps is recommended. This could involve defining more robust name-matching rules between the datasets, as well as improving coordinates-based methods for greater accuracy. Expanding the analysis to other social media platforms, such as TripAdvisor—designed specifically for visitor reviews and recommendations—could also provide richer insights into public perceptions and preferences.

From a cultural perspective, other than complementing traditional safeguarding efforts and improving the efficiency of resource allocation, this methodology promotes a more inclusive, participatory model of heritage management, enhancing both conservation efforts and civic responsibility (Petraroia, 2014).

AKNOWLEDGEMENTS

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.

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

REFERENCES

Abbasi, A., Chen, H. and Salem, A. (2008), "Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums", *ACM Trans. Inf. Syst.*, Association for Computing Machinery, New York, NY, USA, Vol. 26 No. 3, doi: 10.1145/1361684.1361685.

Amato, F., Moscato, V., Picariello, A., Colace, F., De Santo, M., Schreiber, F.A. and Tanca, L. (2017), "Big data meets digital cultural heritage: Design and implementation of SCRABS, a smart context-aware browsing assistant for cultural environments", *Journal on Computing and Cultural Heritage*, Association for Computing Machinery, Vol. 10 No. 1, doi: 10.1145/3012286.

Bai, N., Ducci, M., Mirzikashvili, R., Nourian, P. and Pereira Roders, A. (2023), "MAPPING URBAN HERITAGE IMAGES WITH SOCIAL MEDIA DATA AND ARTIFICIAL INTELLIGENCE, A CASE STUDY IN TESTACCIO, ROME", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLVIII-M–2, pp. 139–146, doi: 10.5194/isprs-archives-XLVIII-M-2-2023-139-2023.

Bai, N., Nourian, P. and Roders, A.P. (2021), "Global citizens and world heritage: social inclusion of online communities in heritage planning", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, International Society for Photogrammetry and Remote Sensing, Vol. 46 No. M-1–2021, pp. 23–30, doi: https://doi.org/10.5194/isprs-archives-XLVI-M-1-2021-23-2021.

Bonci, A., Clini, P., Martin, R., Pirani, M., Quattrini, R. and Raikov, A. (2018), "Collaborative intelligence cyber-physical system for the valorization and re-use of cultural heritage", *Journal of Information Technology in Construction*, Vol. 23 No. 1, pp. 305–323.

Bugalia, N., Tarani, V., Kedia, J. and Gadekar, H. (2022), "Machine Learning-Based Automated Classification of Worker-Reported Safety Reports in Construction", *Journal of Information Technology in Construction*, Vol. 27 No. September, pp. 926–950, doi: 10.36680/j.itcon.2022.045.

Chen, C.-C. and Chang, C.-C. (2024), "Evaluating Public Library Services in Taiwan through User-Generated Content: Analyzing Google Maps Reviews", *Electronics*, Vol. 13 No. 12, doi: 10.3390/electronics13122393.

Chowdhury, M., Hosseini, M.R., Martek, I., Edwards, D.J. and Wang, J. (2021), "The effectiveness of web-based technology platforms in facilitating construction project collaboration: A qualitative analysis of 1,152 user reviews", *Journal of Information Technology in Construction*, Vol. 26 No. March, pp. 953–973, doi: 10.36680/J.ITCON.2021.051.

Council of Europe. (2005), "Council of Europe Framework Convention on the Value of Cultural Heritage for Society (CETS No. 199)", p. 30.

D.Lgs. 42/2004 (Bb.Cc.Aa.). (2004), *Gazzetta Ufficiale*.

D'orazio, M., Bernardini, G. and Di Giuseppe, E. (2023), "Improving Cultural Heritage conservation: LSTM neural networks to effectively processing end-user's maintenance requests", *Vitruvio*, Vol. 8, pp. 18–29, doi: 10.4995/vitruvioijats.2023.18811.

D'Orazio, M., Bernardini, G. and Di Giuseppe, E. (2024), "Influence of Pre-Processing Methods on the Automatic Priority Prediction of Native-Language End-Users' Maintenance Requests Through Machine Learning

Methods", *Journal of Information Technology in Construction*, Vol. 29, pp. 99–116, doi: 10.36680/j.itcon.2024.006.

D'Orazio, M., Di Giuseppe, E. and Bernardini, G. (2022), "Automatic detection of maintenance requests: Comparison of Human Manual Annotation and Sentiment Analysis techniques", *Automation in Construction*, Vol. 134, p. 104068, doi: 10.1016/j.autcon.2021.104068.

Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A.Y.A., Gelbukh, A. and Zhou, Q. (2016), "Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques", *Cognitive Computation*, Springer US, Vol. 8 No. 4, pp. 757–771, doi: 10.1007/s12559-016-9415-7.

Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2019), "BERT: Pre-training of deep bidirectional transformers for language understanding", *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, Vol. 1 No. Mlm, pp. 4171–4186.

Explosion. (2024), "it_core_news_sm", *Github*, available at: https://spacy.io/models/it (accessed 14 October 2024).

Fang, X. and Zhan, J. (2015), "Sentiment analysis using product review data", *Journal of Big Data*, Journal of Big Data, Vol. 2 No. 1, doi: 10.1186/s40537-015-0015-2.

Foroughi, M., de Andrade, B. and Pereira Roders, A. (2023), "Capturing public voices: The role of social media in heritage management", *Habitat International*, Vol. 142, p. 102934, doi: 10.1016/j.habitatint.2023.102934.

Google. (n.d.). "Remove reviews from your Business Profile on Google", *Support.Google*, available at: https://support.google.com/business/answer/4596773?hl=en&co=GENIE.Platform%3DAndroid (accessed 8 August 2024).

Grcheva, O. and Oktay Vehbi, B. (2021), "From Public Participation to Co-Creation in the Cultural Heritage Management Decision-Making Process", *Sustainability*, Vol. 13 No. 16, p. 9321, doi: 10.3390/su13169321.

Hugging Face. (n.d.). "Huggingface.Co", available at: https://huggingface.co/docs/transformers (accessed 12 August 2024).

Hutto, C. and Gilbert, E. (2014), "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text", *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 8 No. 1, pp. 216–225, doi: 10.1609/icwsm.v8i1.14550.

ICCD. (2015), *Normativa A - Architettura 3.00 - Norme Di Compilazione*.

ICCD. (2023), "Registro enti schedatori al 31.12.2023", ICCD.

ICCD. (n.d.). "Catalogo generale dei Beni Culturali", available at: https://www.catalogo.beniculturali.it/ (accessed 16 April 2024a).

ICCD. (n.d.). "SIGECweb", available at: https://www.sigecweb.beniculturali.it/it.iccd.sigec.axweb.Main/ (accessed 26 April 2024b).

ICCD. (n.d.). "ICCD home", available at: http://www.iccd.beniculturali.it/ (accessed 26 April 2024c).

ICCD. (n.d.). "Per catalogare - Diventa ente schedatore", *ICCD*, available at: http://www.iccd.beniculturali.it/it/per-catalogare/diventa-ente-schedatore (accessed 16 April 2024d).

ICCD. (n.d.). "Standard catalografici", *ICCD*, available at: http://www.iccd.beniculturali.it/it/standard-catalografici (accessed 26 April 2024e).

ICCD. (n.d.). "Distribuzione in formato json dei dati delle schede catalografiche 'Scheda Architettura' della regione Marche", available at: https://dati.beniculturali.it/lodview/resource/Distribuzione/arco_CatalogueRecordA_marche-JSON_LD.html (accessed 5 August 2024f).

ICOMOS. (1987), "Charter for the Conservation of Historic Towns and Urban Areas (Washington Charter 1987)".

ICOMOS. (2008), "Charter for the interpretation and presentation of cultural heritage sites", *16th General Assembly of ICOMOS, Québec (Canada), on 4 October 2008*.

ICOMOS, U. (2011), "The Valletta principles for the safeguarding and management of historic cities, towns and urban areas", *Principles, as Adopted by the 17th ICOMOS General Assembly*.

Jacoby, R., Bou-zeid, E., Pigliautile, I. and Laura, A. (2024), "Science of the Total Environment Crowdsourced data as a strategic approach to include the human dimension in outdoor environmental quality assessments", *Science of the Total Environment*, Elsevier B.V., Vol. 953 No. September, p. 175925, doi: 10.1016/j.scitotenv.2024.175925.

Jain, S.M. (2022), "BERT", *Introduction to Transformers for NLP: With the Hugging Face Library and Models to Solve Problems*, Apress, Berkeley, CA, pp. 37–49, doi: 10.1007/978-1-4842-8844-3_3.

Kioussi, A., Karoglou, M., Labropoulos, K., Bakolas, A. and Moropoulou, A. (2013), "Integrated documentation protocols enabling decision making in cultural heritage protection", *Journal of Cultural Heritage*, Elsevier Masson SAS, Vol. 14 No. 3 SUPPL, pp. e141–e146, doi: 10.1016/j.culher.2013.01.007.

Letsmarche. (n.d.). "Città UNESCO nelle Marche", *Let's Marche*, available at: https://letsmarche.it/-/città-unesco-nelle-marche-1#:~:text=La visita di Urbino%2C Pesaro,è davvero un itinerario imperdibile. (accessed 5 August 2024).

Li, J., Krishnamurthy, S., Pereira Roders, A. and van Wesemael, P. (2020a), "Informing or consulting? Exploring community participation within urban heritage management in China", *Habitat International*, Vol. 105 No. October, doi: 10.1016/j.habitatint.2020.102268.

Li, J., Krishnamurthy, S., Pereira Roders, A. and van Wesemael, P. (2020b), "State-of-the-practice: Assessing community participation within Chinese cultural World Heritage properties", *Habitat International*, Vol. 96, doi: @article{sites2008icomos, title={The ICOMOS charter for the interpretation and presentation of cultural heritage sites}, author={SITES, OF CULTURAL HERITAGE}, journal={International Journal of Cultural Property}, volume={15}, number={4}, pages={377--383}, year={2008} }.

Li, J., Krishnamurthy, S., Pereira Roders, A. and van Wesemael, P. (2020c), "Community participation in cultural heritage management: A systematic literature review comparing Chinese and international practices", *Cities*, Vol. 96 No. September 2019, doi: 10.1016/j.cities.2019.102476.

Li, L., Mao, Y., Wang, Y. and Ma, Z. (2022), "How has airport service quality changed in the context of COVID-19: A data-driven crowdsourcing approach based on sentiment analysis", *Journal of Air Transport Management*, Elsevier Ltd, Vol. 105, doi: 10.1016/j.jairtraman.2022.102298.

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

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., *et al.* (2019), "RoBERTa: A Robustly Optimized BERT Pretraining Approach", No. 1.

Lorenzini, E., Calzati, V. and Giudici, P. (2011), "Territorial brands for tourism development: A statistical analysis on the Marche region", *Annals of Tourism Research*, Elsevier Ltd, Vol. 38 No. 2, pp. 540–560, doi: https://letsmarche.it/-/citt%C3%A0-unesco-nelle-marche-1#:~:text=La%20visita%20di%20Urbino%2C%20Pesaro,%C3%A8%20davvero%20un%20itinerario%20imperdibile.

Loyola, M. (2018), "Big data in building design: A review", *Journal of Information Technology in Construction*, Vol. 23 No. November, pp. 259–284.

Mancinelli, M.L. (2018), "Gli standard catalografici dell'Istituto Centrale per il Catalogo e la Documentazione", *Le Voci, Le Opere e Le Cose. La Catalogazione Dei Beni Culturali Demoetnoantropologici*, pp. 279–302.

Marche, R. (n.d.). "Patrimonio culturale", *Regione Marche*, available at: https://www.regione.marche.it/Regione-Utile/Cultura/Patrimonio-culturale.

Masciotta, M.G., Morais, M.J., Ramos, L.F., Oliveira, D. V., Sánchez-Aparicio, L.J. and González-Aguilera, D. (2021), "A Digital-based Integrated Methodology for the Preventive Conservation of Cultural Heritage:

The Experience of HeritageCare Project", *International Journal of Architectural Heritage*, Taylor & Francis, Vol. 15 No. 6, pp. 844–863, doi: 10.1080/15583058.2019.1668985.

Mathayomchan, B. and Taecharungroj, V. (2020), "'How was your meal?' Examining customer experience using Google maps reviews", *International Journal of Hospitality Management*, Pergamon, Vol. 90, p. 102641, doi: 10.1016/J.IJHM.2020.102641.

Medhat, W., Hassan, A. and Korashy, H. (2014), "Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal*, Elsevier, Vol. 5 No. 4, pp. 1093–1113, doi: 10.1016/J.ASEJ.2014.04.011.

Mendes, M.P., Torres-González, M., Valença, J. and Silva, A. (2022), "The Maintenance of Monuments as the Main Trigger to Negative Feelings in Tourists", *Buildings*, MDPI, Vol. 12 No. 12, doi: 10.3390/buildings12122153.

MiC, DGABAP, ICR, I. (n.d.). "Vincoli in Rete", available at: http://vincoliinrete.beniculturali.it/VincoliInRete/vir/utente/login (accessed 16 April 2024).

MiC; ICR. (n.d.). "Carta del Rischio", available at: http://www.cartadelrischio.beniculturali.it/webgis/ (accessed 5 August 2024).

MiC. (n.d.). "Monitoring", *MiC-Ufficio UNESCO*, available at: https://unesco.cultura.gov.it/en/monitoring/ (accessed 30 July 2024a).

MiC. (n.d.). "Centro storico di Urbino", *Miniestero Dell Cultura-Ufficio UNESCO*, available at: https://unesco.cultura.gov.it/projects/centro-storico-di-urbino/ (accessed 5 August 2024b).

MiC. (n.d.). "MiC", available at: https://cultura.gov.it/ (accessed 5 August 2024c).

MiC. (n.d.). "organizzazione", *MiC*, available at: https://cultura.gov.it/organizzazione (accessed 5 August 2024d).

Miller, S.E. and Murray, E.J. (2018), "Heritage Monitoring Scouts: Engaging the Public to Monitor Sites at Risk Across Florida", *Conservation and Management of Archaeological Sites*, Routledge, Vol. 20 No. 4, pp. 234–260, doi: 10.1080/13505033.2018.1516455.

Moioli, R., Van Balen, K. and Vandesande, A. (2014), "Costing the built environment: towards a policy of strong sustainable development. Quale sostenibilità per il restauro", *International Conference "Scienza e Beni Culturali" Proceedings, Arcadia Ricerche, Venezia*, pp. 457–467.

Monteiro, V., Painho, M. and Vaz, E. (2015), "Is the heritage really important? A theoretical framework for heritage reputation using citizen sensing", *Habitat International*, Vol. 45, pp. 156–162, doi: 10.1016/j.habitatint.2014.06.022.

Mountstephens, J. and Quen, M.T.Z. (2023), "Mixed-Language Sentiment Analysis on Malaysian Social Media Using Translated VADER and Normalisation Heuristics", *Lecture Notes in Networks and Systems*, pp. 185–196, doi: 10.1007/978-981-19-9379-4_15.

Murphy, R. (2018), "Comparison of local review sites: Which platform is growing the fastest", *Brightlocal*, available at: https://www.brightlocal.com/research/comparison-of-local-review-sites/# (accessed 8 August 2024).

Nazir, A., Rao, Y., Wu, L. and Sun, L. (2022), "Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey", *IEEE Transactions on Affective Computing*, IEEE, Vol. 13 No. 2, pp. 845–863, doi: 10.1109/TAFFC.2020.2970399.

nlptown. (2020), "bert-base-multilingual-uncased-sentiment", *Hugging Face*, available at: https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment (accessed 6 July 2024).

Petraroia, P. (2014), "La valorizzazione come dimensione relazionale della tutela", *Il Diritto Dell'arte. La Protezione Del Patrimonio Artistico*, Skira Milano.

Picone, R. (2012), "Il rudere architettonico nella storia del restauro", *CONFRONTI*, pp. 27–40.

Pourkeyvan, A., Safa, R. and Sorourkhah, A. (2024), "Harnessing the Power of Hugging Face Transformers for Predicting Mental Health Disorders in Social Networks", *IEEE Access*, IEEE, Vol. 12 No. January, pp. 28025–28035, doi: 10.1109/ACCESS.2024.3366653.

Presidenza del Consiglio dei Ministri. (2021), *DPCM Del 24 Giugno 2021 N123-Regolamento Concernente Modifiche Al Regolamento Di Organizzazione Del Ministero Della Cultura, Degli Uffici Di Diretta Collaborazione Del Ministro e Dell'Organismo Indipendente Di Valutazione Della Performance*.

Quagliarini, E., Gianangeli, A., D'Orazio, M., Gregorini, B., Osimani, A., Aquilanti, L. and Clementi, F. (2019), "Effect of temperature and relative humidity on algae biofouling on different fired brick surfaces", *Construction and Building Materials*, Elsevier Ltd, Vol. 199, pp. 396–405, doi: 10.1016/j.conbuildmat.2018.12.023.

Ravichandiran, S. (2021), *Getting Started with Google BERT: Build and Train State-of-the-Art Natural Language Processing Models Using BERT*, Packt Publishing Ltd.

Reluis. (n.d.). "Terremoto Centro Italia 2016", *Reluis*, available at: https://www.reluis.it/supporto-alle-emergenze-sismiche/terremoto-centro-italia-2016/ (accessed 21 August 2024).

Rosetti, I., Bertrand Cabral, C., Pereira Roders, A., Jacobs, M. and Albuquerque, R. (2022), "Heritage and Sustainability: Regulating Participation", *Sustainability*, Vol. 14 No. 3, doi: 10.3390/su14031674.

Rosin, U., Bonazzi, M., Casarin, F. and others. (2022), "eWom and sentiment analysis to support decision processes within maritime heritage museum networks. The case of Arca Adriatica", *Proceedings of 16th International AIMAC Conference on Arts and Cultural Management*.

Rossi, M. and Bournas, D. (2023), "Structural Health Monitoring and Management of Cultural Heritage Structures: A State-of-the-Art Review", *Applied Sciences (Switzerland)*, Vol. 13 No. 11, doi: 10.3390/app13116450.

Sanoff, H. (1999), *The Valletta Principles for the Safeguarding and Management of Historic Cities, Towns and Urban Areas*, John Wiley \& Sons.

Sergiacomi, C., Vuletić, D., Paletto, A., Barbierato, E. and Fagarazzi, C. (2022), "Exploring National Park Visitors' Judgements from Social Media: The Case Study of Plitvice Lakes National Park", *Forests*, Vol. 13 No. 5, p. 717, doi: 10.3390/f13050717.

"Spacy.Io". (n.d.). , available at: https://spacy.io/ (accessed 8 July 2024).

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

Della Torre, S. (2023), "L'idea di Coevoluzione messa in pratica", *Intrecci-International Journal of Architectural Conservation and Restoration*, No. 3, pp. 4–17.

Torres-Gonzáles, M., Valença, J., Santos, B.O., Silva, A. and Mendes, M.P. (2023), "StainView: A Fast and Reliable Method for Mapping Stains in Facades Using Image Classification in HSV and CIELab Colour Space", *Remote Sensing*, Vol. 15 No. 11, doi: 10.3390/rs15112895.

UAB Webrobots. (n.d.). "Instant Data Scraper", *Chromewebstore*, available at: https://chromewebstore.google.com/detail/instant-data-scraper/ofaokhiedipichpaobibbnahnkdoiiah (accessed 8 August 2024a).

UAB Webrobots. (n.d.). "Instant Data Scraper", *Webrobots.Io*, available at: https://webrobots.io/instantdata/ (accessed 8 August 2024b).

UNESCO. (2011), "Recommendation on the Historic Urban Landscape", *Records of the General Conference - 31st Session*, Vol. 1 No. November, p. 170.

UNESCO. (n.d.). "Italy", *UNESCO*, available at: https://whc.unesco.org/en/statesparties/it (accessed 4 September 2024a).

UNESCO. (n.d.). "UNESCO creative cities", *UNESCO*, available at: https://www.unesco.org/en/creative-cities (accessed 5 August 2024b).

Valdivia, A., Martínez-Cámara, E., Chaturvedi, I., Luzón, M.V., Cambria, E., Ong, Y.-S. and Herrera, F. (2020), "What do people think about this monument? Understanding negative reviews via deep learning, clustering and descriptive rules", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11 No. 1, pp. 39–52, doi: 10.1007/s12652-018-1150-3.

Valença, J., Mendes, M.P., Jouen, N., Morin, K., Olivo, N. and Silva, A. (2024), "Emotional Maintenance: A Digital Model to Support Maintenance Decisions in Buildings' Coatings", *Engineering Proceedings*, No. November, p. 53, doi: 10.3390/iocbd2023-16488.

Vandesande, A. and Van Balen, K. (2018), "Preventive conservation applied to built heritage: a working definition and influencing factors", *Innovative Built Heritage Models*, CRC Press, pp. 63–72.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., *et al.* (2017), "Attention is All you Need", in Guyon, I., Luxburg, U. Von, Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S. and Garnett, R. (Eds.), *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc.

Wankhade, M., Rao, A.C.S. and Kulkarni, C. (2022), *A Survey on Sentiment Analysis Methods, Applications, and Challenges*, *Artificial Intelligence Review*, Vol. 55, Springer Netherlands, doi: 10.1007/s10462-022-10144-1.

World Economic Forum. (2020), "These are the countries with the most UNESCO World Heritage sites", *World Economic Forum*, available at: https://www.weforum.org/agenda/2020/12/unesco-world-heritage-countries-history-archaeology-landmarks-italy-china/ (accessed 4 September 2024).

World Heritage Committee. (2015), "Policy Document for the Integration of a Sustainable Development Perspective to the Processes of the World Heritage Convention", *United Nations Educational, ScientifiQc, and Cultural Organization: Geneva, Switzerland*.

Zhang, W., Li, X., Deng, Y., Bing, L. and Lam, W. (2023), "A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges", *IEEE Transactions on Knowledge and Data Engineering*, IEEE, Vol. 35 No. 11, pp. 11019–11038, doi: 10.1109/TKDE.2022.3230975.