

Review

Social Media Platforms and Computational Approaches for Analyzing Visitor Experience in Museums and Cultural Heritage Sites: A Literature Review

Georgios Yfantidis and Panagiotis D. Michailidis * 

Department of Balkan, Slavic & Oriental Studies, University of Macedonia, Egnatia 156, 54636 Thessaloniki, Greece; yfantidisgeorge@uom.edu.gr

* Correspondence: pmichailidis@uom.edu.gr

Abstract

Social media has become an important tool for understanding visitor experiences in museums and cultural heritage sites. This literature review identifies, organizes, and thematically synthesizes existing studies on museum visitor experience based on social media data. It examines 41 studies retrieved from Scopus and Web of Science and published between 2017 and 2026. Furthermore, the review examines the selected studies across six dimensions: social media platforms, types of user-generated data, the role of digital interactions, the museums and cultural heritage sites studied, the analytical methodologies applied, and the main findings on visitor experience. The findings indicate that TripAdvisor is the most frequently used platform for collecting textual reviews and star ratings, whereas Instagram and Flickr are mainly used for visual and spatial data. Most studies rely on computational methods, often combined with quantitative techniques, while qualitative approaches are used less frequently. The identified methods include content analysis, statistical analysis, sentiment analysis, topic modeling, machine learning, image analysis, and spatial analysis. Across the reviewed studies, visitor experience is examined as a multidimensional phenomenon encompassing emotions, service quality, authenticity, historical connection, aesthetics, education, and social participation. Finally, the review identifies recurring themes across the dimensions and synthesizes them into broader research streams. These are brought together in an integrative synthesis framework that organizes existing research, highlights research gaps, and outlines directions for future studies.

Keywords: visitor experience; social media; computational methods; natural language processing; machine learning; cultural heritage; digital humanities



Academic Editors: Phivos Mylonas, Christos Troussas, Akrivi Krouska, Manolis Wallace and Maria Kouri

Received: 20 July 2026

Revised: 17 August 2026

Accepted: 21 August 2026

Published: 24 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Cultural heritage sites and museums play an important role in promoting knowledge and understanding of art and history and preserving cultural identity [1,2]. During their interactions with these cultural sites and museums, visitors express their overall perceptions, satisfaction and emotional experiences. This is referred to as the visitor experience. Visitor experience can be understood as a multidimensional process that emerges from visitors' interactions with museum collections, cultural spaces, services, staff and other visitors. It may include cognitive, emotional, aesthetic, social and functional dimensions [1]. The quality of the visitor experience has become increasingly important, as museums seek not only to present collections but also to create meaningful, engaging, and memorable experiences for their audiences [1,3–5].

In recent years, the widespread use of social media platforms has allowed visitors to communicate, evaluate and share experiences online, as well as interact with cultural spaces digitally. Social media platforms such as TripAdvisor, Google Reviews, Facebook, Instagram and Twitter/X provide large amounts of user-generated content (UGC), including textual reviews, ratings, photographs, comments, and interaction data [6,7]. However, social media platforms should not be regarded simply as neutral channels through which visitor experiences are recorded. The concept of social media logic emphasizes that platforms mediate communication and interaction through mechanisms such as programmability, popularity, connectivity and datafication [8]. Although social media data do not capture the visitor experience in its entirety, they provide access to the aspects of the experience that users choose to express and share publicly. They therefore represent publicly expressed and platform-mediated aspects of visitor experience rather than the experience as a whole. Such data can provide valuable insights into visitors' perceptions, emotions, satisfaction, and engagement with museums and cultural heritage sites [2,9]. Furthermore, these data can highlight the factors associated with these experiences, which can contribute both to the management and promotion of cultural heritage and to the strengthening of the relationship between museums and the public. Therefore, the analysis of museum visitors' experiences through social media platforms has emerged as a contemporary and rapidly evolving field of research [10–13].

However, the large volume and unstructured nature of social media data create analytical challenges that cannot easily be addressed through manual approaches alone. Social media analytics has consequently developed as an interdisciplinary field concerned with the discovery, collection, preparation, and analysis of social media data using a range of analytical approaches [14]. Within research on museums and cultural heritage, computational methods such as Natural Language Processing (NLP), sentiment analysis, topic modeling, social network analysis, machine learning and deep learning allow researchers to understand and identify patterns in visitors' expressed perceptions, emotions and behavior [9,15,16]. Using these methods, researchers and museums can gain insights into publicly expressed aspects of visitor experience, which may inform interpretation and communication, management and audience-engagement practices [2,17].

Although previous studies have applied computational methods to social media data to analyze visitor experience, the literature remains fragmented across museum studies, cultural heritage research and data analytics. This fragmentation makes it difficult to understand how different platforms, types of user-generated data and computational methods have been used to examine visitor experience in museums and cultural heritage sites. Recent research also highlights the analysis of social media data as an important yet still underexplored area for studying museum visitor experiences. It also suggests the need for further investigation of how social media shape visitor experiences and how computational methods can be applied to user-generated data [4,10,11,18].

Previous reviews provide important foundations for this research but address different aspects of the field. Vassiliadis and Belenioti [12], based on a review of 54 studies, examined the broader role of social media in museums, organizing the literature around opportunities for museum experience and communication, enhancement of learning, patterns of social media use, and barriers to social media integration. More recently, Nigatu et al. [4] conducted a bibliometric analysis of 407 Scopus-indexed journal articles on the broader field of museum visitor experience, mapping its development and intellectual and social structure and identifying future research directions. Within that review, social media and big data were identified as an area requiring further research. Although these reviews provide important foundations, a gap remains in critically synthesizing social media-based visitor-experience research across platforms, forms of user-generated data, analytical ap-

proaches, cultural heritage contexts, and dimensions of visitor experience. The present review therefore focuses specifically on this intersection in museums and cultural heritage sites. In this context, particular emphasis is placed on the computational approaches used to analyze social media-based visitor experience and their relationship with platforms, user-generated data, cultural contexts and visitor-experience dimensions. Table 1 summarizes the differences in focus, methodology and contribution between previous reviews and the present study.

Table 1. Comparison with previous reviews.

Study/Year	Focus	Methodology	Contribution
Vassiliadis & Belenioti [12]	Social media in museums: museum experience and communication, learning, patterns of social media use, and barriers to social media integration	Literature review	Synthesizes the role and use of social media in museums
Nigatu et al. [4]	Broad museum visitor-experience research	Bibliometric analysis	Maps the broader museum visitor-experience research field and identifies emerging research directions
Our study	Social media-based visitor experience in museums and cultural heritage, with emphasis on platforms, user-generated data, computational methods and experience dimensions	Narrative literature review	Provides an integrated synthesis of social media-based visitor-experience research and highlights key research gaps and future directions

Therefore, this review aims to synthesize the existing evidence on social media-based visitor experience in museums and cultural heritage sites. It examines the platforms, forms of user-generated data, analytical approaches, cultural contexts and visitor-experience findings reported in the literature. The review then compares findings across these dimensions to identify recurring themes, organize them into broader research streams, and highlight methodological and empirical gaps that can inform future research.

The review is therefore guided by the following main research question:

RQ: What major themes emerge from the literature on social media-based visitor experience in museums and cultural heritage sites, and how can these themes be organized into research streams?

Finally, the recurring themes identified by comparing findings across the analytical dimensions are brought together in an integrative synthesis framework for social media-based visitor-experience research. The framework organizes recurring patterns in the reviewed literature into broader analytical research streams. It is intended to synthesize existing knowledge, highlight comparatively underexplored areas and guide future research rather than to represent an empirically validated explanatory model.

2. Review Methodology

To answer the research question of this review, this study adopts a narrative literature review approach. The review followed a structured search, selection and synthesis process to enhance transparency while retaining the interpretive and integrative purpose of a narrative review [19]. The review was organized into three phases: (1) Literature search and scope, (2) Selection of relevant studies and (3) Thematic synthesis. These phases were

used to identify relevant studies, organize the literature and highlight major methodological trends, findings and research gaps.

2.1. Literature Search and Scope

To identify relevant studies, searches were conducted in the two scientific databases, Scopus and Web of Science. These databases were selected because of their broad coverage of peer-reviewed articles on cultural heritage, visitor experiences and social media analysis.

To identify relevant articles, the literature search used combinations of keywords related to social media platforms and visitor experiences in museums and cultural heritage institutions. Based on the literature, the related terms used for social media and digital content environments included: social media, online platform, TripAdvisor, Google my business, Facebook, user-generated content, Instagram, social media analytics, online review and Flickr. Correspondingly, the keywords used for visitor experiences in museums and cultural heritage included: museum, cultural heritage, digital heritage, visitor interactions, museum experience, and visitor experience. Both generic terms and specific platform names were included in order to capture broader social-media research while retaining relevance to commonly studied digital environments. The following search query was submitted to Scopus and Web of Science using the Boolean operators OR and AND. The final search query which was applied to the title, abstract and keywords fields of the databases was as follows:

“social media” OR “online platform” OR “user-generated content” OR “tripadvisor” OR “google my business” OR “facebook” OR “instagram” OR “flickr” OR “social media analytic*” OR “online review”) AND (museum* OR “cultural heritage” OR “digital heritage” OR “visitor interactions” OR “museum experience*” OR “visitor experience*”)

The search covered publications from 2017 to 2026 and was limited to English-language publications. The final search was conducted in May 2026. The selected period was intended to focus the review on recent developments in social media-based visitor-experience research. The restriction to Scopus and Web of Science and to English-language publications provides a clearly defined evidence base but may underrepresent studies published in other databases, languages and regional contexts, including research involving locally dominant social media platforms and heritage contexts. This potential geographic and linguistic limitation was therefore considered when interpreting the patterns identified in the review.

2.2. Selection of Relevant Studies

All articles identified through the database searches were combined into a single dataset, and duplicates were removed using Microsoft Excel. The remaining articles were screened on the basis of their titles and abstracts. Potentially eligible publications were then assessed at full-text level to confirm their relevance to the scope of the review.

Studies were included when they (i) focused on museums or cultural heritage sites, (ii) examined visitor experience, perceptions, emotions, satisfaction or digital engagement, (iii) used data derived from social media platforms, (iv) applied qualitative, quantitative, or computational analytical approaches, (v) were published as journal articles or conference papers, (vi) were written in English and (vii) were published between 2017 and 2026.

The study-selection process is summarized in Figure 1 using a PRISMA-type flow diagram [20], which reports the numbers of records identified, duplicates removed, records screened, full texts assessed and publications excluded at each stage. The diagram is included to improve transparency in reporting the identification and selection of studies. This selection process resulted in a final corpus of 41 studies, including 36 journal articles and 5 conference papers. These studies formed the basis of the thematic analysis presented in the following sections.

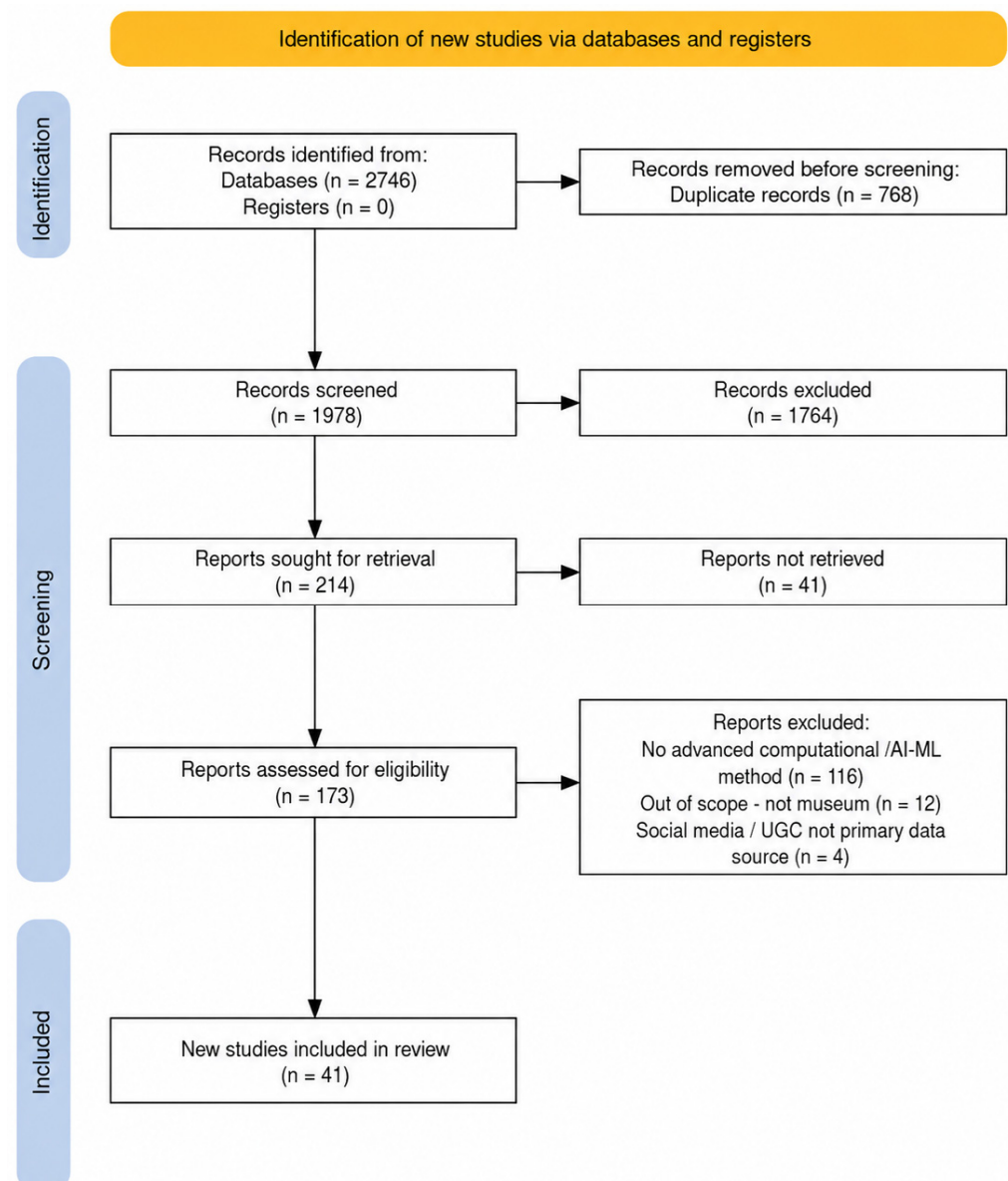


Figure 1. Workflow of the review methodology.

2.3. Thematic Synthesis

The 41 selected articles were examined in depth and organized for analysis according to six dimensions: the social media platforms used for data collection, the types of user-generated data examined, the role of digital interactions in shaping visitor experience, the types of cultural sites studied, the analytical approaches used and the main findings related to visitor experience. These dimensions served as an organizing structure for comparing the reviewed studies. The synthesis was informed by established principles of thematic analysis, which emphasize the systematic identification, comparison and iterative refinement of patterns within a body of material [21].

For each dimension, the categories were developed and refined iteratively through repeated reading and comparison of the selected studies. The purpose of the classification was to identify patterns, methodological trends, underexplored areas and connections between platforms, data types, analytical approaches and visitor-experience dimensions. Following this initial organization, findings were compared across the six dimensions to identify broader recurring themes relevant to the research question. These cross-dimensional themes were subsequently synthesized into broader research streams, which are developed in the

discussion section. To enhance the trustworthiness and consistency of the categorization, coding was carried out independently by both authors, and discrepancies were resolved through discussion.

It should be noted that some studies addressed more than one platform, data type, cultural setting, or computational method; therefore, they were assigned to more than one category where necessary. The categories are presented in the summary tables and/or figures, and the reported frequencies represent the number of studies associated with each category. Because studies could be assigned to multiple categories, these frequencies are not expected to sum to the total number of studies reviewed. The frequencies are used descriptively to indicate patterns in the reviewed literature and should not be interpreted as measures of methodological quality or conceptual importance. Given the narrative and interpretive nature of the review, no formal quality-appraisal tool was applied; instead, methodological differences across studies were considered descriptively during the synthesis.

3. Findings by Analytical Dimension

This section presents the findings from the 41 selected studies according to the six analytical dimensions described above. These findings provide the basis for the subsequent comparison and synthesis of recurring themes presented in the Discussion.

3.1. Social Media Platforms

This section examines digital platforms as a key source of data through which visitors express their opinions and experiences of visits to museums and other cultural heritage sites. Figure 2 shows the frequency of use of various digital platforms for collecting data related to visitors' experiences in museums and cultural sites. The analysis of the reviewed literature identified TripAdvisor as the dominant platform, with seventeen (17) studies, indicating its primary role in recording and expressing visitors' experiences. These articles are presented by platform in Table 2.

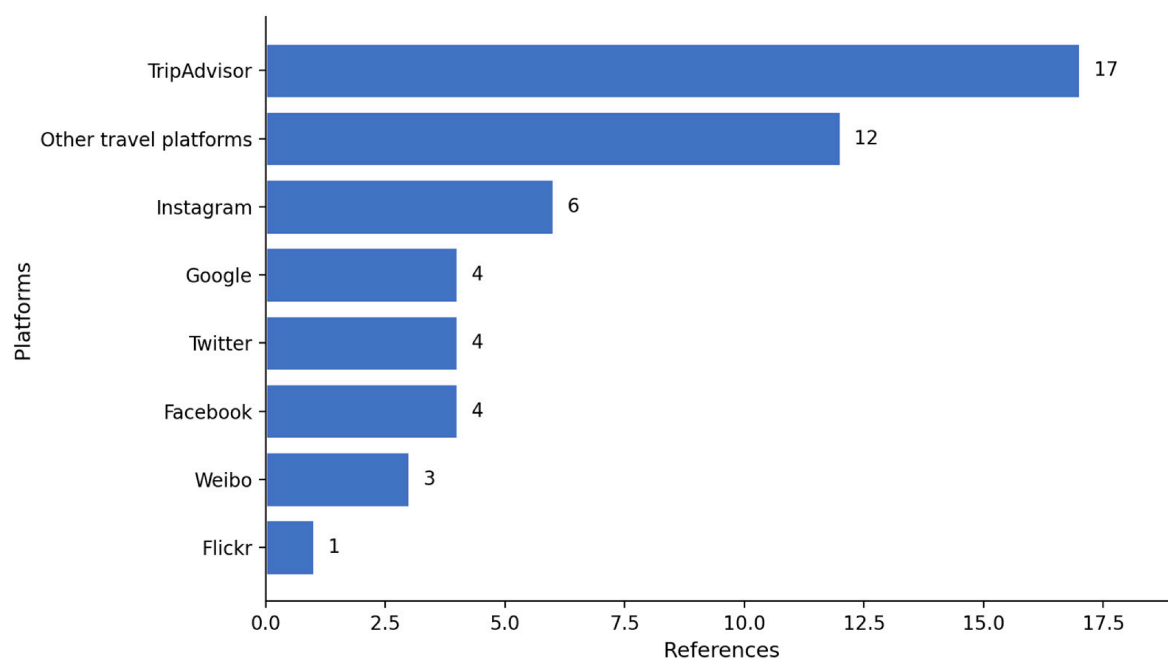


Figure 2. Number of studies by platform.

Table 2. Articles by platform.

Platform	References
TripAdvisor	[2,3,9,22–35]
Other travel platforms	[11,17,18,25,36–43]
Instagram	[28,44–48]
Google Reviews	[5,10,15,28]
Twitter/X	[27,28,49,50]
Facebook	[27,28,50,51]
Weibo	[16,41,42]
Flickr	[52]

TripAdvisor was chosen as a platform for collecting data on museum visitor experiences, as it offers a large volume of authentic user-generated reviews that capture visitors' impressions, emotions and evaluations. The fact that this platform is dominated by textual reviews posted without the mediation of the museums themselves is a crucial factor in enhancing the credibility of these opinions. Essentially, the main form of content used is written text, although the platform allows the use of images and videos by users [23]. The frequent use of TripAdvisor is linked to the availability of a large volume of user-generated reviews, which are often accompanied by star ratings and user metadata.

Instagram follows appearing in six (6) studies. These articles are listed in Table 2. Instagram was used as a data collection platform because it is an inherently visual medium in which users share photos, accompanied by captions and hashtags that frame their experience. The emphasis on the aesthetic presence of the spaces and the artworks, as well as the participatory dimension (interaction during the visit, sharing after the visit) makes Instagram a useful platform for studying how visitors experience, interpret and share their visits. The data collected consist mainly of images accompanied by short texts, such as captions and hashtags. Although the platform also supports videos, most studies focus on analyzing visual content in relation to aesthetics, emotions and popularity [47,48].

Twitter/X was examined in four (4) studies [27,28,49,50]. Twitter/X was used as a data collection tool because it allows researchers to capture visitors' immediate, spontaneous and public reactions and discourse concerning cultural events. Through the use of text, hashtags and images, users capture emotions, impressions and evaluations, which makes it suitable for the analysis of the aesthetic and emotional experience of the audience [27].

Flickr appears in one (1) study [52]. It was used as a platform for collecting data related to visitors' experiences at cultural sites, as it offers geographically localized and time-stamped images accompanied by descriptive texts and comments.

Facebook was examined in four (4) studies [27,28,50,51]. It was used as a data collection platform as it is one of the most popular social media platforms and encourages users to share their experiences through photos, short descriptions and hashtags. The nature of the platform, which supports the audiovisual representation of the experiences, makes it particularly suitable for documenting visitor participation in cultural spaces [51].

A considerable number of studies also relied on other travel and review platforms, such as Xiaohongshu, Dianping and Ctrip, which together appeared in twelve (12) studies. Google Reviews appears with four (4) studies [5,10,15,28] and Weibo with three (3) studies [16,41,42]. The choice of platform is linked to the aspect of visitor experience being examined. Therefore, the social media platform is not merely a technical source of data, but also influences which aspects of the visitor experience become visible. Overall, the diversity of platforms identified highlights the multidimensional nature of visitor experience in

museums. Specifically, the studies included in this review were strongly concentrated on text-based review platforms, while purely visual and multimodal social-media environments received comparatively less attention.

It should be noted that several methods are used to collect data from social media platforms. One common approach is the automatic extraction of data from review platforms such as TripAdvisor and Google Reviews through web-scraping tools. Representative examples of studies adopting this approach include Drivas et al. [10] and Pinheiro et al. [31]. This practice is particularly widespread because it allows researchers to collect large volumes of visitor comments for subsequent analysis. Another approach involves the use of an Application Programming Interface (API), which allows researchers to send a query to a server and retrieve specific data. For example, Gerrard et al. [49] used the Twitter API to collect tweets around museum events and hashtags, while Vu et al. [52] used the Flickr API to collect geotagged photos, GPS coordinates, timestamps and user metadata. These API-based approaches provide structured access to data, but are constrained by platforms' terms of use, the availability of API endpoints, and changes in access policies. In addition to these automated data collection techniques, some studies relied on manual data collection and selection, as shown by Agostino et al. [22] and Rhee et al. [47].

3.2. Types of User-Generated Data

This section examines the data produced by social media users to record their experiences. These data are classified as textual data, visual data, star ratings and several other types of data, as shown in Figure 3. These data contribute to understanding visitors' expectations and perceptions.

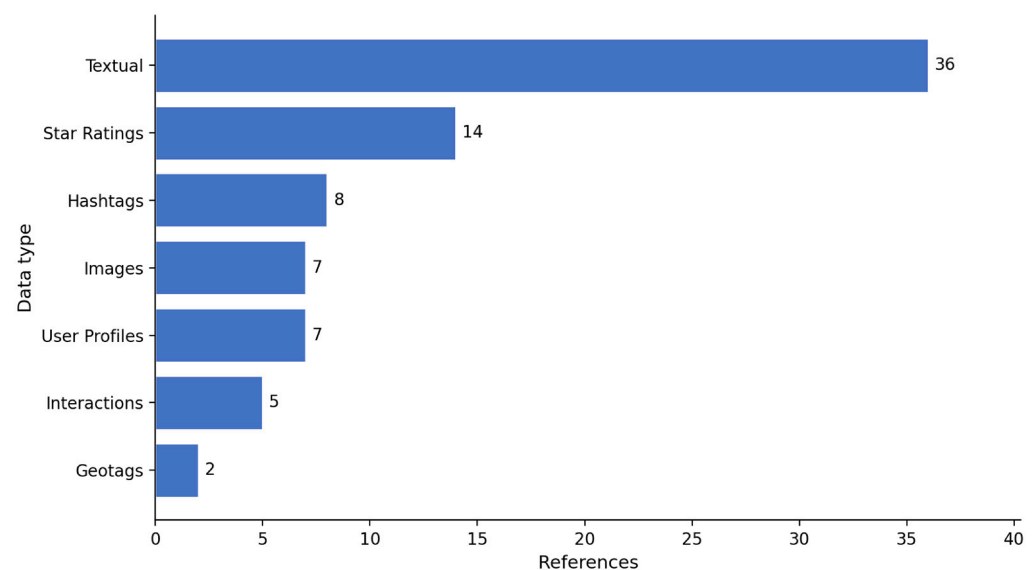


Figure 3. Number of studies by data type.

Table 3 presents the articles assigned by data type. Textual reviews on social platforms are an important source of data for understanding visitors' experiences of artworks and their popularity. The analysis of posts combined with emotion recognition through machine learning algorithms, allows researchers to understand how visitors experience and communicate art in the digital space [11].

Table 3. Articles categorized by data type.

Data Type	References
Textual	[2,3,5,9–11,15–18,22–44,46,49,50]
Star Ratings	[2,5,10,22–26,30–32,34,35,37]
Hashtags	[11,16,44–47,49,50]
Images	[44,46–48,50–52]
User Profiles	[3,9,11,16,29,38,45]
Interactions (follows, likes, etc.)	[10,11,16,45,51]
Geotags	[47,52]

Star ratings, typically measured on a scale of 1–5, were examined in 14 studies. Star ratings serve as key indicators of overall visitor experience and provide a basis for classifying reviews into positive and negative categories. Reviews capture visitors' perceptions of the quality of services offered. Star ratings allow for the quantitative investigation of visitor satisfaction and can highlight important factors for improving the museum experience [23].

Images provide complementary visual data and were examined in seven (7) studies. Social media images are a valuable source of data for capturing and analyzing public perceptions and visual representations of cultural heritage, as they combine visual content with metadata. Their analysis using artificial intelligence techniques, such as computer vision, can highlight both expected and less prominent elements of cultural interest [47].

Hashtags were examined in eight (8) studies. Social networks can function as complementary sources of information for monuments and museums. Hashtags emerge as a key means of organizing and retrieving content, although spelling variations in monument names may create ambiguities [46].

User profiles were examined in seven (7) studies. User profiles have clear analytical value, as information such as membership level provides additional context about users and may enhance the reliability of the findings [45].

Five (5) studies examined user interactions on social media, such as likes and shares. Visitor interactions on social media are not limited to a simple recording of the visit, but actively contribute to shaping the museum experience. Specifically, these interactions show not only what the original visitor thought, but also how other users reacted to the shared content. As a result, they serve as a means of personal expression and social communication, offering valuable knowledge for the design of more participatory and enriched museum experiences [51].

Furthermore, geotags were examined in two studies. Geotags are an important source of data for analyzing visitor experience in museums and cultural sites. The use of these data provides researchers with an opportunity to connect the visual experience with the physical environment, identify patterns of visitation and explore the distribution of visitor interest in the space. Furthermore, they provide spatial information by recording the location of an experience without relying solely on the visitor's subjective narrative [52].

Several studies combine more than one form of data. Multimodal data allow for a more complex and dynamic understanding of cultural objects. The multimodal approach offers a multidimensional understanding of cultural heritage [44].

Overall, the diversity of data types indicates that social media platforms are valuable tools for understanding cultural experience beyond the capabilities of traditional methods. Furthermore, textual reviews were the most commonly used data type because they are readily available and are frequently accompanied by star ratings. However, images, spatial and video data can enhance the ability to identify patterns in visitor experience.

3.3. Role of Digital Interactions

Table 4 presents the articles based on the role of digital interactions.

Table 4. Articles categorized by the role of digital interactions.

Role of Digital Interactions	Description	References
UGC as direct experiential feedback	Studies that use user-generated content (reviews, comments, ratings) as direct feedback to assess visitor experience, satisfaction and perceptions.	[2,3,5,9,10,15,17,18,22–35,37–43]
UGC as input for brand/destination image and management	Studies where digital interactions inform brand and destination image, reputation and marketing or management strategies.	[11,16,18,36,45,51]
UGC as collective memory and social meaning-making	Studies in which digital interaction functions as a form of collective memory, shared and affective experience and identity construction.	[44,46,49,50]
UGC as visual self-presentation and digital popularity	Studies on visual sharing practices (photographic participation, ‘instagrammability’) and digital popularity.	[47,48,52]

Digital interactions are a critical source of data for understanding visitor experiences. Online reviews, ratings and posts provide unsolicited visitor feedback that can reveal recurring perceptions and concerns. Nevertheless, this content is shaped by user self-selection, platform characteristics and patterns of online participation. Through analyses of data provided by users on social media, patterns of visitation, areas of interest and emotional trends emerge, allowing researchers and managers to map visitor preferences and improve the management of cultural spaces [22].

In a similar context, digital interactions can significantly shape the public image of cultural destinations and visitor behavior. Positive reviews and narratives may enhance visitors’ intentions to revisit or recommend, while negative comments may discourage potential visitors. Furthermore, digital interactions can reveal the socio-political and cultural dimensions of the experience by highlighting aspects such as learning, authenticity and emotional engagement with cultural heritage. In this sense, digital interactions are not simply interpreted as reflections of visitor experience, but also function as an active field for the strategic promotion and sustainable management of cultural heritage sites and museums [16,36].

Equally important is the emerging field of collective memory construction and social experience shaped by visitors’ digital interactions in cultural spaces. Through texts, images and metadata, visitors construct narratives that transcend individual experience and help shape perceptions of cultural heritage at a global level. The posting and dissemination of this content do not simply capture personal impressions, but also function as a collective record of memory, contributing to the formation of the public image of museums and places of cultural interest [44,50].

Finally, it is worth noting that digital interactions contribute significantly to shaping the aesthetic experience. Through comments, likes and visual patterns, emotions such as joy, awe or attraction emerge, influencing the reception and popularity of works of art. At the same time, because digital platforms allow researchers to record responses to exhibitions and artistic activities, they can function as indicators of public interest and participation.

Although this dimension initially arises from personal viewing, it subsequently extends to social interaction and the co-creation of meaning by visitor communities [47,48].

3.4. Types of Museums and Cultural Heritage Sites

Table 5 presents the articles according to the type of cultural site examined.

Table 5. Articles categorized by the type of cultural site.

Type of Cultural Site	References
General and Multi-museum Studies	[2,9,10,15,22–24,26,36,38,49,52]
Art Museums and Galleries (including digital/immersive art spaces)	[5,11,42,45–48,50,51]
History, National, Thematic and Specialized Museums	[3,18,25], [28] *, [33–35,37], [40] *, [41,43]
Architectural/Built Heritage and Landscapes	[16] *, [17], [31] *, [32,39]
Archaeological and Religious Sites	[27,29,30] *
Memorial and Dark-Heritage/Memory Sites	[44] *
* UNESCO World Heritage Site (cross-cutting tag)	Eight studies examine a UNESCO World Heritage Site: [16,27–31,40,44]

According to the reviewed literature, the largest group of studies does not focus on a single museum type but examines museums more broadly through city-wide or multi-museum datasets. Art museums and cultural exhibition spaces form the most prominent specific category, followed by historical, national and thematic museums; architectural and built heritage sites and landscapes; archaeological and religious sites; and, more rarely, memorial or dark-heritage sites.

At the same time, a notable strand of research focuses on cultural sites of specific cities and regions to document local specificities and cultural tourism development strategies. Typical examples include museums in London [23], museums in St. Petersburg (Russia) [26], the leading tourist attractions in Antalya (Turkey) [24], and Italian state museums [9,22]. These examples show how local traditions and global cultural products are intertwined in contemporary research.

One of the most prominent areas of research is art museums and cultural exhibition spaces, which constitute one of the largest specific museum categories. The studies analyze large, internationally renowned museums, such as the Louvre, the Museum of Modern Art in New York and the British Museum [11]. Alongside these institutions, specialized and thematic exhibitions are also examined, such as the World Press Photo photography exhibition [46] and instagrammable art exhibitions [47,48]. Research on hybrid and digital museums, which combine traditional museum functions with modern forms of entertainment, also occupies a special place and includes studies of immersive art museums [5] and digital art museums [42]. This focus reflects continued interest in the relationship between culture, art and the public.

The literature suggests interest in a wide range of cultural sites that vary in both type and geographical distribution. In this context, sites associated with historical memory and dark tourism have been examined, such as the Auschwitz memorial, highlighting the sensitivity of this research to issues of historical memory [44]. At the same time, several studies focus on monuments and World Cultural Heritage Sites, such as the Terracotta Army [30], the Forbidden City and other heritage sites in Beijing [16,40], the Grand Canal [28], the

historic center of Coimbra [31] and the Borobudur and Prambanan temples [29], thus integrating the research into the field of global cultural management.

Finally, the literature includes studies that focus on more specialized or alternative cultural spaces. This category includes thematic museums, such as mountain museums [3], and historic residences repurposed for cultural use, such as the Ganxi Former Residence in Nanjing [39], as well as cultural and natural sites beyond the conventional museum, such as the Majorelle Garden in Marrakech [32]. These examples highlight new forms of access to and engagement with cultural heritage.

3.5. Analytical and Methodological Approaches

This section explores qualitative and quantitative methods, as well as computational techniques, used to draw conclusions from visitor data. In this review, these approaches are presented together within a unified classification scheme (Table 6), in which qualitative and quantitative methods function as complementary layers to the dominant computational analysis.

Table 6. Methods and analytical approaches (qualitative, quantitative and computational).

Method	References
A. Computational text analysis	
Lexical analysis and text mining	[3,11,15–17,22,24–26,32,34,35,37–39,41–43,46,47,49,50]
Sentiment & emotion analysis	[3,10,11,15,17,18,24,25,27–29,31–33,35–40,42,43]
Topic modeling	[2,3,5,9–11,15,18,22–24,27,28,30,32–35,38,41,43,46]
Semantic network analysis	[17,36,37,42]
B. Computational visual & multimodal	
Computational visual & image analysis	[44,47,48,51]
Multimodal/multi-data-type analysis	[44,51]
C. Network, engagement & spatial-temporal	
Engagement analysis	[10,11,50,51]
Social network & diffusion analysis	[16,45]
Spatial, geotag-based & temporal analysis	[38,52]
D. Statistical/quantitative analysis	
Statistical, explanatory & predictive analysis	[2,5,10,11,15–18,22–26,28–32,34–36,38,39,43,50–52]
E. Qualitative analysis (complementary)	
Qualitative textual interpretation	[9,28,34,41,44,46,49]
Qualitative visual interpretation	[44,46,48,52]

3.5.1. Computational Methods

Computational methods are prominent among the studies included in the review for the analysis of visitor experience, as they allow the processing of large-scale and complex social media data sets. In the relevant literature, computational methods are commonly applied through a set of recurring analytical processes, including data collection, preprocessing, feature extraction, model selection, validation and interpretation. As shown in Table 6, the main computational methods presented in the literature include computational text analysis, image and multimodal data analysis, spatial analysis and social network analysis.

Across the reviewed studies, the computational analysis of visitor experience commonly involves processes such as data acquisition, data preprocessing, feature extraction,

computational modelling, validation and interpretation, although their implementation varies according to the data and analytical task.

Data acquisition is usually conducted through web scraping, platform APIs or manually collected datasets. Preprocessing involves cleaning textual data, removing duplicates and irrelevant records, handling missing values, and transforming unstructured data into formats suitable for computational analysis. For textual data, preprocessing may also include tokenization, stop-word removal, stemming or lemmatization, depending on the analytical approach. For visual data, preprocessing includes removing duplicate and irrelevant images [47], and preparing them for subsequent analysis through resizing, normalization or other standard image-processing procedures. Feature extraction then converts the preprocessed data into meaningful numerical representations that can be used by computational models. In textual analysis, this may involve techniques such as TF-IDF, word embeddings or contextual embeddings, whereas visual analysis typically relies on feature vectors extracted through computer vision or deep learning models. Depending on the analytical objective, dimensionality reduction techniques (such as UMAP) may subsequently be applied to facilitate visualization, clustering or other analyses [44]. These representations are then used in natural language processing, machine learning, topic modeling, image analysis or spatial analysis methods to identify patterns related to visitor experience. Finally, computational models are evaluated through validation procedures appropriate to the analytical task, such as classification performance metrics, topic coherence measures or expert assessment. However, several studies do not explicitly report validation procedures or performance metrics, which limits methodological reproducibility.

The review identified five broad categories of analytical approaches, with computational text analysis being the dominant category. This trend reflects the extensive use of textual reviews, especially from platforms such as TripAdvisor, where visitors express their evaluations, emotions and perceptions through written comments and ratings. Computational text analysis includes techniques such as lexical and text-mining approaches, semantic network analysis, topic modeling and sentiment analysis. In particular, lexical and text mining approaches are the most common techniques and constitute the initial stage of the computational framework. This phase is used to extract frequent words, keywords, TF-IDF terms, n-grams, word clouds and other surface-level textual patterns from online reviews, comments, captions and posts. For example, some studies used lexical analysis as an exploratory step to identify the aspects of museum experience most frequently mentioned by visitors, including exhibitions, history, culture, tickets, queues, children, staff, accessibility and facilities [17,35,37,38,43]. Moreover, these word frequencies enable the data to be utilized for either statistical processing or specific techniques such as sentiment analysis, topic modeling and semantic network analysis.

Sentiment Analysis (SA) is one of the most widely used applications of Natural Language Processing (NLP). It is used to determine the polarity (positive, negative, neutral) of user opinions and is applied to textual reviews from various platforms. For example, TripAdvisor is used to analyze museum reviews, while platforms such as Instagram and Twitter are used to study the captions, hashtags and comments associated with the aesthetic experience of artworks. Three main computational approaches are used to implement sentiment analysis: lexicon-based methods, Machine Learning (ML), and Deep Learning (DL). Lexicon-based methods use rule-based models such as VADER, TextBlob, SnowNLP and ROST-CM. They are useful for social media texts because they are transparent, easy to implement and do not require training data [3,32,38,42]. In these cases, labelling is usually automatic: reviews or comments are assigned positive, negative or neutral polarity according to word-level sentiment scores and predefined thresholds. However, they may perform less effectively in cases involving irony, cultural nuance, domain-specific vocabulary or

multilingual reviews. Machine learning methods require pre-labeled datasets or rely on star ratings as proxy labels as well as train/test splits to classify reviews into sentiment categories. These approaches are implemented with algorithms such as Support Vector Machine (SVM), Naïve Bayes, Random Forest and Logistic Regression, which recognize sentiments and patterns [28,29,35]. At the same time, deep learning approaches utilize more sophisticated models such as BERT, which provides stronger contextual representations and records improved prediction performance compared to previous methods [15,39]. However, deep learning models require larger datasets, greater computational resources and clearer validation procedures. The level of validation varies considerably across the literature. Some studies provide explicit validation through accuracy, precision, recall, F1-score, confusion matrices or model comparison [15,29,31,35,43]. In contrast, many lexicon-based studies report sentiment distributions without strong external validation or manually annotated ground truth. This means that sentiment labels are often treated as computational outputs rather than independently verified interpretations of visitor emotions.

In addition to sentiment analysis, topic modeling algorithms are applied to automatically recognize patterns, group and categorize large amounts of data. This process is often implemented with algorithms such as Latent Dirichlet Allocation (LDA), which is an unsupervised model. Topic modeling helps to discover latent structures in large text collections and detect topics related to visitor experiences at World Heritage Sites [30]. Initial studies [22,23] use topic modeling to identify topics in museum reviews, while later studies apply more advanced approaches such as Structural Topic Modeling (STM) and BERTopic. Specifically, structural topic modeling is used to relate topics to ratings and time [2], whereas BERTopic or transformer-supported topic extraction are used to identify more fine-grained visitor-experience dimensions [5,33,43]. Nevertheless, topic coherence, parameter selection and the interpretability of generated topics should be explicitly reported, as these factors influence the validity of the findings.

Semantic network analysis is used to capture relationships and interconnections between words, topics and concepts [37]. Unlike topic modeling methods, which group words into latent themes, semantic network analysis focuses on the structure of associations among concepts, making it useful for understanding how visitors connect ideas such as authenticity, memory, learning, emotion, atmosphere and satisfaction. Semantic network-based indicators are used to examine cultural reputation [36], while semantic or co-occurrence networks are used to show how visitors connect cultural value, services, emotions, and spatial aspects of the museum experience [17,37,42]. In the literature, this approach appears less frequently than sentiment analysis or topic modeling, indicating a methodological gap and a possible direction for future research.

Less frequently, the literature also includes the categories of image and multimodal data analysis, social network analysis, engagement analysis, and spatial analysis. Image and multimodal data analysis includes computer vision and visual analysis techniques which, in combination with metadata, contribute to the interpretation of the visual and emotional dimensions of the cultural experience. For instance, deep-learning computer vision has been used to automatically analyse and tag the visual content of visitors' images, recognising objects, artworks and architectural elements [47,48]. Tools such as the Google Vision API facilitate the automatic assignment of labels to large sets of images [51]. In multimodal approaches, images and their captions are converted into feature vectors and grouped through dimensionality reduction and clustering techniques in order to identify thematic clusters of visual representations [44]. These studies show that computational image analysis can reveal aspects of visitor experience that are not captured by text alone, such as visual attention, photo-taking behavior and self-presentation. However, many studies still analyze images separately from textual reviews, ratings or geotags. This gap suggests

a future research direction toward multimodal models that integrate text, image, spatial metadata and interaction metrics for a holistic understanding of the visitor experience.

Social network analysis focuses on the structure and circulation of digital content. For example, Chang et al. [45] examine how museums are connected through hashtag-based networks to identify clusters of museums, shared thematic associations and patterns of digital visibility among art museums. Furthermore, Wang et al. [16] analyze social media discussions on Weibo, focusing on how information circulates through posts, reposts, comments and influential users to identify opinion leaders and diffusion patterns. On the other hand, the computational analysis of engagement metrics, such as likes and comments, has been used to assess the popularity of posts and exhibitions and to identify which cultural objects attract attention or stimulate digital engagement [10,11,51].

Finally, spatial analysis utilizes geographic and temporal metadata, allowing researchers to investigate visitation patterns, map mobility patterns and identify geographic concentration of activities. Vu et al. [52] use geotagged Flickr photos and spatial clustering to examine visitation patterns, while Jiang et al. [38] apply GIS-based techniques to explore the spatial distribution of museum experience in China. These studies show that geospatial metadata links visitor experience to specific places, routes and spatial concentrations of cultural interest. However, spatial analysis remains underdeveloped compared to text-based methods, despite its potential for heritage management and site interpretation.

3.5.2. Quantitative Methods

Quantitative methods are important research tools, as they allow the systematic recording and analysis of large datasets, providing possibilities for generalization and comparative evaluation. The different techniques and applications identified in the literature are presented in Table 6. The literature review shows that quantitative methods are used for the analysis and interpretation of large and multidimensional data. Within these, basic statistical analysis occupies a prominent position, including descriptive statistics to summarize visitor evaluations, correlation and regression analyses, and tests of statistical significance, offering an overall picture of trends and relationships between variables. For example, Drivas et al. [10] applied correlation analysis to show that longer and more emotionally rich reviews are associated with stronger expressed sentiment, illustrating how basic statistics can reveal significant relationships between review characteristics and visitors' expressed experience.

At the same time, regression methods are applied to predict and assess effects such as the impact of the physical, cultural and emotional characteristics of a cultural heritage site on visitor satisfaction, with the aim of identifying the factors that contribute most to a positive experience. For example, Shin et al. [5] combine topic modeling and regression analysis within a Kano-model framework to identify which museum attributes act as satisfiers or dissatisfiers, clarifying the factors that most strongly shape visitor satisfaction and dissatisfaction. Another clear example is the study by Huo et al. [2], which combines topic modeling with factor analysis techniques and structural equation modeling (SEM) to examine how museum attributes influence personal experience and satisfaction. In addition, a number of studies such as Hariyono et al. [29], Pinheiro et al. [31], Ricciardi and Manisera [15] and Llewellyn et al. [50] report validation or classification metrics such as accuracy, precision, recall, F1-score or AUC to evaluate machine learning and deep learning models. Overall, quantitative approaches provide statistical evidence and support systematic comparison, often complementing computational and qualitative methods to enhance research validity.

3.5.3. Qualitative Methods

Qualitative methods play an important role in understanding visitors' experiences, attitudes and perceptions in depth. In the reviewed literature, however, they appear in a limited number of studies, mostly as a complementary, interpretive layer within computational approaches, and are presented in Table 6.

The reviewed literature shows that qualitative methods are generally used in combination with computational techniques rather than autonomously. Among them, content analysis is the most commonly used qualitative text analysis approach. It often involves manual coding or an inductive analytical process and is applied to the interpretation of textual data such as comments and reviews. It is frequently combined with text mining techniques [34]. Thematic analysis is also used to interpret patterns and themes across various datasets, typically in combination with computational topic modeling [28,41]. In addition, some studies apply supervised, top-down content analysis in which predefined categories guide the coding [22].

By contrast, qualitative visual interpretation is concentrated in image-based studies [44,46,48,52]. These studies use manual or interpretive visual analysis to understand how visitors represent museum experience through images, photographs, poses and other visual practices. Overall, within this corpus qualitative methods are used mainly as a complementary, interpretive layer that helps capture meanings, experiences and social processes from data produced in physical and digital cultural environments.

3.6. Findings on Visitor Experience

This subsection synthesizes the visitor-experience dimensions identified across the reviewed studies. As shown in Table 7, these dimensions include emotional responses, service quality and infrastructure, the role of social media in cultural experience, aesthetics, authenticity, learning, social participation and related experiential outcomes.

Table 7. Articles categorized by visitor-experience dimensions.

Dimensions	References
Emotional Dimension	[2,3,5,10,11,17,18,24–35,37–43,49,51]
Service Quality & Infrastructure	[2,3,5,10,17,18,22–24,26,30,32–35,37–43]
Role of Social Media in Cultural Experience	[11,16,18,25,27,36–38,42,44–47,49–52]
Visual Representation & Aesthetics	[5,11,18,39,42,44,46–48,50–52]
Educational Dimension	[2,3,16–18,28,33–35,37–41,43,44]
Visitor Profiles & Differences	[3,9,11,15,16,29,38,45,52]
Social Participation & Public Interpretation	[11,16,27,44–51]
Authenticity, Memory & Historical Connection	[16–18,30,37–39,41,43,44]

The analysis of visitors' experiences can reveal the emotions and values that emerge during their exposure to cultural heritage. Visitors frequently express strong emotional responses such as awe, admiration, enjoyment and surprise which intensify the aesthetic and experiential dimensions of the visit [31,41].

Accordingly, positive evaluations are associated with factors such as authenticity, historical connection and atmosphere; for example, visitors to the Terracotta Army express amazement and pride in its historical significance [30], while attention to historical artifacts and cultural symbols evokes positive emotions such as awe and beauty in heritage muse-

ums [41]. Conversely, negative impressions are attributed to high prices, cleanliness problems, long queues, overcrowding and poor organization; reviews of the Palace Museum, for instance, highlight tourist flow, ticketing, toilets and cost as sources of dissatisfaction [40], and overcrowding emerges as a recurring problem in landmark museums [34,41]. The quality of services and staff interaction also largely determines the level of satisfaction [26].

Social media plays an important role in the construction of the experience, allowing visitors to represent and extend their experiences through images, captions and hashtags [47]. Social media users who visit museums or cultural sites act as co-creators of content, shaping the digital image of the spaces and influencing public perception [11]. The visual content shared by visitors, together with the engagement it generates, also shapes the popularity of exhibitions and cultural spaces [48].

The educational and cognitive dimensions are also important elements in the formation of the experience, as the educational content of museums enhances learning and supports lifelong learning [28]. Finally, the cultural context and the specific characteristics of each space shape the visitor experience, with notable differences between local and non-local or international visitors [3,9,52].

At the same time, participation through social media enhances the aesthetic experience, social interaction and the dissemination of cultural content, with visitors actively co-creating and circulating content about cultural heritage [16,45,46].

Authenticity, historical significance and the state of conservation are key valuation factors at World Heritage Sites and heritage museums, where visitors emphasise the cultural and historical value of the site [17,30,41].

At genocide memorials such as Auschwitz, visitors use social media to attribute historical significance to the site and to emotionally process their experience, participating in the collective construction of memory [44].

Across the reviewed studies, visitor experience is treated as a holistic and multidimensional process in which cognitive, emotional, aesthetic and social components interact. Overall, the thematic synthesis indicates that the reviewed literature is concentrated on text-based platforms, textual data and computational text analysis. On the other hand, visual, spatial, interaction and multimodal approaches are represented less frequently. The findings also show that visitor experience is treated as a multidimensional construct involving emotional, functional, aesthetic, educational, social and heritage-related dimensions.

4. Discussion

A synthesis of the 41 selected studies highlights several recurring themes that emerge across the different dimensions examined in the review. In this section, we discuss these themes and present an integrative synthesis framework that brings together the main patterns in the literature into broader research streams. We also discuss the limitations and gaps identified in the existing literature and outline directions for future research.

4.1. The Predominance of Text-Based Platforms and Data

The reviewed literature shows a clear predominance of TripAdvisor and other review-oriented platforms, accompanied by a strong reliance on textual reviews and ratings. This concentration also influences the aspects of visitor experience that are most often examined. Review-oriented platforms mainly provide direct evaluations from users, making emotions, satisfaction, service quality, and infrastructure particularly visible in the existing literature.

In contrast, visual platforms such as Instagram and Flickr receive much less attention in the literature, despite their close connection to the aesthetic and experiential dimensions of museum visits. Images and hashtags are also examined less frequently than textual reviews. As a result, visual, spatial and other aspects of visitor experience expressed

through non-textual content receive comparatively less attention. This imbalance therefore reflects not only the platforms selected by researchers but also the types of visitor experience that can be examined through the available data.

The choice of digital platform shapes which aspects of visitor experience become visible to researchers, because each platform provides different forms of content, interaction and metadata. TripAdvisor appears dominant in textual reviews and functional service quality, while Instagram emphasizes the visual representation and aesthetic dimension of the visit. Therefore, findings based on social media data reflect the aspects of visitor experience that users choose to express on particular platforms rather than visitor experience in its entirety. In addition, because these data come from visitors who choose to post or review online, they may not represent the wider population of museum and heritage visitors.

4.2. Methodological Diversity and Data–Method Relationships

Regarding the analytical approaches, the findings show a combination of traditional and computational techniques. Computational text analysis (such as topic modeling, sentiment analysis and related NLP techniques) is particularly prominent, while qualitative, visual, spatial and network-based approaches appear less frequently. Statistical and quantitative techniques are also widely used, either on their own or in combination with computational methods. The findings indicate a methodological variety from traditional analytical procedures toward computational and algorithmic approaches for analyzing visitors' experiences.

Furthermore, this pattern is closely related to the types of data most commonly used in the reviewed studies. For example, studies based on review-oriented platforms such as TripAdvisor and Google Reviews, as well as several heritage-sites studies, make extensive use of computational text analysis, while machine learning is mainly used within specific classification, clustering or predictive tasks. This trend is driven by the large volume of data generated on social media and online review platforms. As a result, researchers adopt text-based computational methods, such as sentiment analysis and topic modeling, to extend rather than simply replace traditional approaches. On the other hand, visual, spatial and interaction data have received less methodological attention.

At the same time, recent research shows a growing interest in combining qualitative and computational methods to use not only textual data but also visual social media content. For example, computer vision and deep learning techniques are used for the automatic recognition of objects, artworks and architectural elements in visitors' images of museums and cultural sites, while dimensionality reduction and clustering methods are applied to identify thematic clusters in large image datasets [44,47].

In contrast, specialized approaches remain less common: For example, semantic network analysis, which examines the structural relationships between words and concepts, is used in fewer studies than sentiment analysis and topic modeling. Spatial analysis, social network analysis, multimodal approaches, and the use of video data are even less common. In particular, video-based analysis remains largely underexplored in the reviewed literature. These areas therefore represent important research gaps. The need to manage large volumes of digital data and uncover deeper thematic and emotional patterns is driving a shift from traditional content analysis to computational methods and machine learning.

However, the increasing use of computational methods also raises questions about how their results are interpreted. The outputs of these methods such as sentiment labels or topic distributions, should be treated as algorithmic approximations rather than direct measures of a visitor's actual sentiment. Interpreting them as experiential meaning requires caution and, ideally, validation against human judgment.

The reviewed studies therefore show both the growing importance of computational approaches and the need for greater attention to the validation and interpretation of their results. Combining computational analysis with human coding, qualitative interpretation or other forms of validation could provide a stronger basis for understanding what these outputs indicate about visitor experience.

Overall, the findings show a close relationship between the type of available data and the methods used to analyze them. For example, textual data are mainly analyzed using NLP techniques, whereas visual and spatial data are used much less frequently. This relationship between data and method is one of the recurring patterns identified across the reviewed literature and shows that methodological choices are closely connected to both the platform and the type of data available.

4.3. Visitor-Experience Patterns Across Cultural-Site Types

The review analysis reveals meaningful patterns depending on the type of cultural organization or museum. The largest group consists of general and multi-museum studies, which examine visitor experience across heterogeneous museum samples rather than focusing on a cultural-site type. A small body of research focuses on archaeological and religious sites and, as a cross-cutting group, UNESCO World Heritage Sites, with an emphasis on authenticity, memory, historical connection and service quality. This trend indicates that the visitor experience in these sites is perceived as a synthesis of cultural experience and operational or tourist service.

In art museums, galleries and visual/digital cultural spaces, a strong relationship is observed between the themes of visual representation and the role of social media in cultural experience. This pattern reflects both the visual character of these institutions and the opportunities offered by image-based social media platforms for self-presentation, photographic participation and digital sharing.

In contrast, dark tourism rarely appears in the reviewed literature, indicating that the digital representation of memory at trauma sites does not yet attract the same research attention as monuments with high tourist traffic. Given the limited number of studies, however, this finding is better understood as an underrepresented area in the reviewed literature rather than as evidence of lower relevance. The intersection of memory, trauma and digital interactions remains largely unexplored.

Research interest is shaped by the type of cultural site: at World Heritage Sites, experience is primarily associated with authenticity and service quality, while in art museums, it focuses on public display and visual aesthetics. Visitor satisfaction is understood as a holistic process where the emotional dimension of the experience is closely linked to the quality of the infrastructure, the historical authenticity and the aesthetics of the space. Taken together, these findings point to context-dependent patterns in visitor-experience research. These patterns should not be understood as fixed characteristics of particular site types, since they may also reflect differences in the platforms, data and analytical approaches used across studies.

4.4. An Integrative Synthesis Framework for Social Media-Based Visitor Experience Analysis

Based on the findings of the thematic synthesis, Figure 4 presents an integrative synthesis framework for organizing the main recurring patterns in social media-based visitor-experience research in museums and cultural heritage sites. The framework is not intended to represent a tested statistical model. Instead, it provides a way of organizing the main research streams that emerged from the comparison of the reviewed studies.

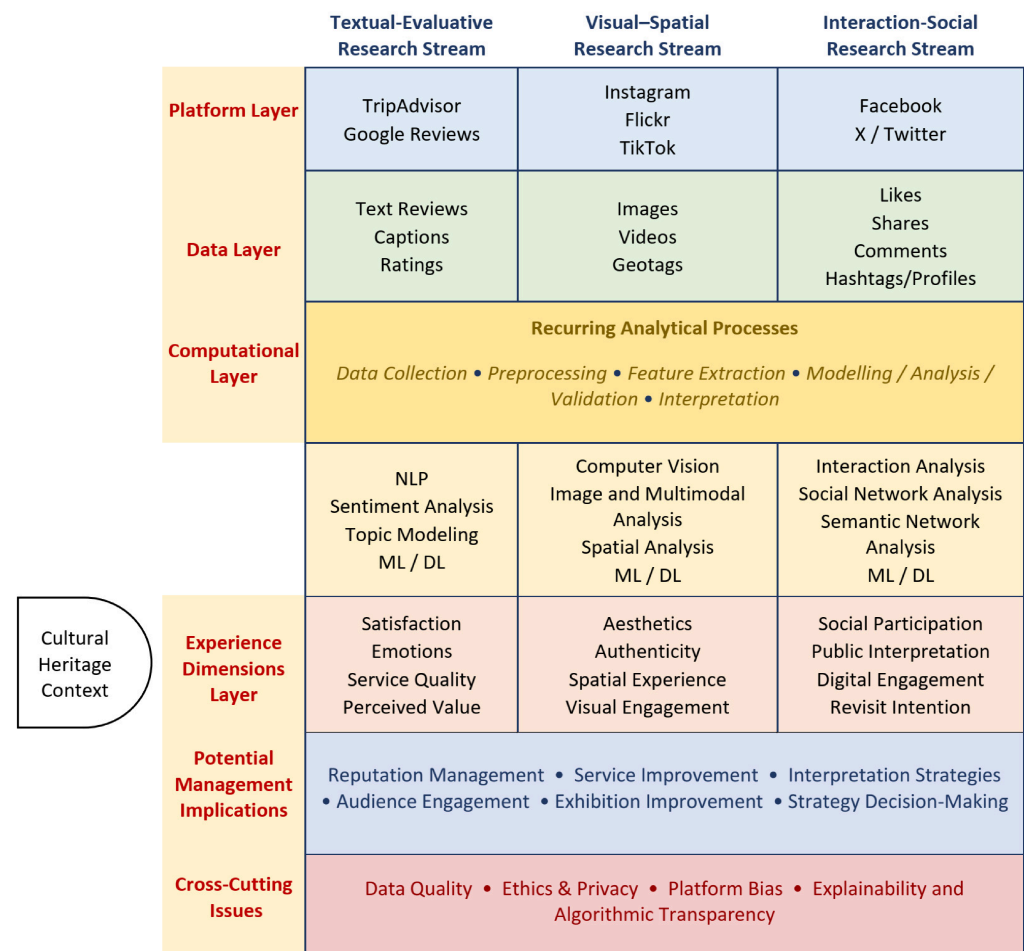


Figure 4. An integrative synthesis framework for social media-based visitor experience research in museums and cultural heritage sites.

Three broader research streams can be identified from recurring combinations of platforms, data types, analytical approaches and dimensions of visitor experience. The textual-evaluative stream, linked to text-based platforms such as TripAdvisor and Google Reviews, uses reviews and ratings to analyze satisfaction, emotions, service quality, and perceived value, and reflects a strong concentration of existing literature. The visual-spatial stream, associated with image-based platforms such as Instagram and Flickr, draws on images and geotags to explore aesthetics and spatial experience, but remains less developed, with video-based platforms such as TikTok and Instagram representing a largely untapped source. The interaction-social stream focuses on likes, comments, shares, and user profiles to capture participatory aspects of visitor experience, yet is still less systematically analyzed, highlighting a gap especially in interaction and social network analysis.

These three streams are not intended to represent completely separate categories. Some studies may include elements from more than one stream, and the boundaries between them may change as platforms, data and analytical methods develop. The streams are mainly used here to organize recurring patterns in the literature and to show areas that have received more or less research attention.

As shown in Figure 4, the computational layer consists of analytical processes that recur across the reviewed studies including data collection, preprocessing, feature extraction, modelling or analysis, validation and interpretation. These processes are not intended to imply that all studies follow a single standardized or linear workflow. Textual data are typically analyzed using NLP, sentiment analysis, topic modeling, text classification, and machine or deep learning techniques. Visual-spatial data are examined using computer

vision, image analysis, multimodal analysis, and spatial analysis. Interaction and social data are often analyzed using interaction analysis, semantic network analysis, and social network analysis.

This distinction helps to illustrate that, while some analytical processes are shared, the computational techniques depend on the type and structure of the data. For example, text reviews and ratings are well suited to sentiment analysis, topic modeling, and classification, while images and geotags require computer vision and spatial analysis. Similarly, likes, comments, hashtags, and user profiles call for interaction or network-based approaches to capture how visitor experiences circulate socially. These recurring combinations help to explain the main methodological patterns observed in the literature, while leaving room for studies that combine different types of data and methods.

The framework also links computational outputs to key dimensions of visitor experience within the cultural heritage context. This context provides an interpretive frame of reference, rather than a particular step in the analysis, because various types of museums and heritage sites can influence how visitation experiences are interpreted and what aspects stand out. For example, World Heritage, archaeological, and religious sites often highlight authenticity, historical connection, and spatial experience, while art museums and cultural institutions tend to emphasize aesthetics, visual engagement, and social media visibility. Similarly, thematic and specialized museums, and dark tourism or memory sites may focus more on learning, emotional engagement, public interpretation, and collective memory. These associations reflect patterns observed in the reviewed corpus and should not be interpreted as stable relationships that necessarily apply across all cultural, linguistic or platform contexts.

Finally, the framework links visitor experience dimensions to potential heritage management outcomes. Social media analytics may inform reputation management, service improvement, interpretation strategies, exhibition improvement, audience engagement, strategic decision-making and sustainable heritage management. However, these outcomes should not be understood as automatic consequences of computational analysis. Instead, they require careful interpretation of findings in relation to the cultural, social and institutional context of each museum or heritage site as well as consideration of the quality and representativeness of the social media data used.

Across the entire framework, several cross-cutting issues should be considered. These include data quality and representativeness, ethics and privacy, platform bias, and explainability and algorithmic transparency. These issues do not form a separate analytical layer; rather, they affect all stages of the process. At the platform level, they influence which visitor voices and experiences become visible. At the level of data, they affect the completeness, reliability and representativeness of user-generated content. On the methodological level, they concern preprocessing decisions, algorithmic bias, validation procedures and the interpretability of computational models. At the visitor experience level, they influence how computational outputs are translated into experiential dimensions such as satisfaction, authenticity, emotion or learning. At the management level, they shape the responsible use of analytics for decision-making and cultural heritage management. These considerations are therefore integral to the interpretation of the three research streams rather than secondary limitations of the analytical process.

To summarize, the framework brings together the main recurring patterns identified across the reviewed literature. It shows the strong concentration of research in the textual-evaluative stream, while visual-spatial and interaction-social research remains comparatively less developed. In this way, the framework provides a synthesis of existing research and helps to identify combinations of platforms, data and methods that have received less attention and may be explored in future studies.

4.5. Limitations of Existing Studies

Despite the increasing interest in the use of social media and computational approaches to analyze visitor experience in museums and cultural heritage sites, the reviewed literature presents several limitations. First, many studies are affected by platform bias, as they rely on data from a single platform or from a limited group of platforms. For example, some studies focus on data from specific platforms such as TripAdvisor [22,23], Google Reviews [10], Instagram [44] and Flickr [52]. While these platforms provide rich user-generated content, they represent only those visitors who choose to post, review or share online. As a result, the findings may reflect platform-specific audiences rather than the full diversity of museum visitors. This should be taken into account when findings based on social media data are used to describe visitor experience more broadly.

A second limitation concerns language and translation issues across the reviewed literature. Several analyses are restricted to one language or to reviews written in English, which may exclude important visitor groups. For example, Saoualih et al. [32] focus on English-language TripAdvisor reviews, while many Chinese studies, such as Gao and Yu [37], Jiang et al. [38], Lan et al. [17], Pan et al. [40] and Xue et al. [43], depend on Chinese-language platforms. Some studies explicitly engage with multilingual data or language-based differences, such as Kirilenko et al. [30], Ricciardi and Manisera [15], Pinheiro et al. [31] and Xu and Shih [35], but multilingual and cross-cultural comparisons remain uneven across the literature. This limits the comparability of results and makes it difficult to distinguish between platform effects, language effects and cultural differences in visitor experience. Language can also affect the results of automated text analysis, particularly sentiment and emotion classification, and therefore requires careful consideration when findings are compared across different linguistic and cultural contexts.

Thirdly, many studies are limited by their short temporal depth. Because most analyses rely on datasets collected at a specific period, the result is a snapshot of visitor experience. Although some studies incorporate time more explicitly, such as Pan et al. [40] through monthly emotional trends and Huo et al. [2] through time-related topic modeling, temporal analysis remains relatively underdeveloped. This limits our understanding of how the way visitors think and feel and how they experience satisfaction in response to events, unfold over time.

Finally, the field is characterized by different computational methods and limited comparability across studies. The reviewed literature uses a diverse set of techniques that include topic modeling, sentiment analysis, statistical modelling, social network analysis, image analysis and machine learning. However, these approaches often use different datasets, platforms, preprocessing procedures and validation strategies. The methodological diversity is a strength but it also makes it difficult to compare findings across studies about visitor experience. Differences in annotation, preprocessing and model-validation procedures may also influence the reported results, making it difficult to determine whether variation across studies reflects visitor experience itself or methodological choices. Therefore, greater transparency in data collection, preprocessing, coding or annotation decisions, model validation and reporting would enhance the comparability and reproducibility of future research.

4.6. Future Research Directions

Despite the rapid growth of platforms such as TikTok and Instagram, research still focuses on text-based data. This highlights the need for the systematic analysis of video as a data source. As the visitor experience becomes more interactive and hybrid, future research should adopt multimodal approaches that combine text, image and sound. The analysis

of the data associated with the dominance of textual data and TripAdvisor confirms that multimodal and cross-platform analyses remain at an early stage of development.

Future research could also examine a wider range of geographic and cultural settings. Alongside highly visible museums and internationally recognized heritage sites, regional museums, memory sites, dark-tourism settings and other less frequently studied cultural institutions could provide useful contexts for comparison. This would help to address some of the geographic and cultural imbalances found in the existing literature.

Another area for future research is cross-platform comparative analysis. Future studies could compare visitor experiences across different social media platforms. They could also examine how platform-specific features influence user-generated content. Such comparisons could also help researchers understand whether observed differences are mainly related to the platform itself or to differences in visitor experience.

Only one of the reviewed studies explicitly adopted an Explainable Artificial Intelligence (XAI) approach through the use of SHAP [43]. As machine learning and deep learning models become more common in visitor experience research, the need for interpretability is increasing. As a result, future work might be directed towards developing explainable models. These models can show how specific features (e.g., words, images or interaction patterns) contribute to sentiment classification and topic identification. Explainability, however, should be used together with appropriate validation so that computational results can be interpreted more reliably.

Future research could also examine the use of large language models and multimodal AI models in visitor-experience analysis. These models may support multilingual analysis, sentiment analysis, topic modeling, aspect extraction, image interpretation and the synthesis of large volumes of user-generated content. However, their outputs should be carefully validated and combined with human interpretation, since they may reproduce platform, language or cultural biases.

A significant research gap in the existing literature on cultural experience is the limited use of Aspect-Based Sentiment Analysis (ABSA). Most studies use traditional sentiment analysis, which may not fully capture mixed user reactions within the same text. For example, visitors may express enthusiasm for the exhibits while also reporting dissatisfaction with crowding. Future research could address this gap by applying ABSA methods. These approaches can link specific sentiments to individual aspects of the visitor experience, providing a more detailed and targeted evaluation of cultural heritage services. Future studies using these methods should also examine how well they perform across different platforms, languages and cultural contexts before using their results to draw conclusions about visitor emotions or evaluations.

Future research could focus on data quality, platform bias and algorithmic bias. It could also examine the ethical issues related to the use of user-generated content, especially when computational methods are used to support decision-making in cultural heritage management. For example, social media data do not represent all visitors equally, since they are shaped by self-selection, language use, moderation policies and recommendation algorithms. For this reason, future research should provide greater transparency regarding data collection procedures, preprocessing decisions, model validation, language filtering, anonymization and ethical compliance.

5. Conclusions

This review synthesized 41 selected studies examining visitor experience in museums and cultural heritage sites through social media platforms. The findings indicate that this research field is increasingly oriented toward the use of computational methods to manage large volumes of user-generated data. Within the reviewed literature, TripAdvisor

emerges as the dominant source of data on the visitor experience, while textual reviews and star ratings are the most commonly used data types. This predominance, however, should be interpreted in relation to the platform-specific and self-selected nature of social media data and does not represent visitor experience in its entirety. The analysis highlights the need to strengthen visual and network-based approaches. It further emphasizes the importance of extending research to less-studied types of cultural heritage beyond globally recognized monuments.

The review also shows that the selected studies approach visitor experience as a multi-dimensional phenomenon, encompassing emotions, service quality, authenticity, historical connection, aesthetics, education and social participation. Methodologically, the prominence of computational approaches, including sentiment analysis, topic modeling, social network analysis, image analysis, spatial analysis and machine learning in the reviewed literature reflects the increasing analytical importance of large-scale user-generated data.

In summary, this review adds to the literature by organizing platforms, data types, digital interactions, cultural site types, methodological approaches and experience outcomes. Comparison across these analytical dimensions reveals three recurring themes: the concentration of research on text-based platforms and data, the close relationship between data types and analytical methods, and the context-dependent variation in visitor-experience patterns across cultural sites. These themes are synthesized into three broader research streams: textual-evaluative, visual-spatial and interaction-social. The textual-evaluative stream is the most developed and is mainly based on online reviews, ratings, sentiment analysis, topic modeling and service evaluation. The visual-spatial stream remains less developed but highlights the importance of images, geotags, aesthetics and spatial experience. The interaction-social stream is also less systematically examined, although it is important for understanding digital participation, social sharing, public interpretation and engagement. These recurring patterns are brought together in an integrative synthesis framework, which is intended to organize the existing literature rather than to function as an empirically validated explanatory model. It therefore provides a basis for future research in computational visitor-experience analysis.

However, we identified several limitations regarding the review design itself. First, the choice of keywords used in the search strategy may affect the completeness of the review by excluding studies that use other relevant terms. In addition, the review was limited to studies retrieved from Scopus and Web of Science, which may have excluded relevant publications indexed in other databases such as Google Scholar. The restriction to English-language publications may also have underrepresented research published in other languages and studies focusing on locally dominant social media platforms or regional cultural contexts. These limitations should be considered when interpreting the patterns identified in the reviewed literature.

Finally, although the literature covers a wide geographical range, examining monuments and museums in different countries, coverage remains geographically uneven, and several regional contexts are underrepresented, with Greek cultural institutions representing one such example. Future research should therefore examine the role of social media data in less-studied regional contexts, including Mediterranean, Balkan and Greek cultural heritage settings, to capture local characteristics and inform museum management.

Author Contributions: G.Y. and P.D.M., methodology: G.Y. and P.D.M., formal analysis and visualization: G.Y., writing—original draft: G.Y., review and editing: G.Y. and P.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Falk, J.H.; Dierking, L.D. *The Museum Experience Revisited*; Routledge: Abingdon, UK, 2016.
2. Huo, H.; Shen, K.; Han, C.; Yang, M. Measuring the relationship between museum attributes and visitors: An application of topic model on museum online reviews. *PLoS ONE* **2024**, *19*, e0304901. [\[CrossRef\]](#)
3. Candrea, A.N.; Ciobanu, E.; Nechita, F.; Brătucu, G.; Coman, E.; Şchiopu, C.; Alexandrescu, M.B. How do visitors to mountain museums think? a cross-country perspective on the sentiments decoded from tripadvisor reviews. *Electronics* **2025**, *14*, 1637. [\[CrossRef\]](#)
4. Nigatu, T.F.; Trupp, A.; Teh, P.Y. A bibliometric analysis of museum visitors' experiences research. *Heritage* **2024**, *7*, 5495–5520. [\[CrossRef\]](#)
5. Shin, S.; Ko, S.-B.; Kang, J. Sustainable tourist satisfaction in art museums: Identifying attributes that enhance visitor experience for sustainable cultural management. *Sustainability* **2026**, *18*, 1367. [\[CrossRef\]](#)
6. Kaplan, A.M.; Haenlein, M. Users of the world, unite! The challenges and opportunities of social media. *Bus. Horiz.* **2010**, *53*, 59–68. [\[CrossRef\]](#)
7. Zeng, B.; Gerritsen, R. What do we know about social media in tourism? A review. *Tour. Manag. Perspect.* **2014**, *10*, 27–36. [\[CrossRef\]](#)
8. Van Dijck, J.; Poell, T. Understanding Social Media Logic. *Media Commun.* **2013**, *1*, 2–14. [\[CrossRef\]](#)
9. Riva, P.; Agostino, D. Latent dimensions of museum experience: Assessing cross-cultural perspectives of visitors from Tripadvisor Reviews. *Mus. Manag. Curatorship* **2022**, *37*, 616–640. [\[CrossRef\]](#)
10. Drivas, I.C.; Vraimaki, E.; Lazaridis, N. I can't get no satisfaction? from reviews to actionable insights: Text data analytics for utilizing online feedback. *Digital* **2025**, *5*, 35. [\[CrossRef\]](#)
11. Sun, Y.; Plaza, B. Clustering museum visitors on Xiaohongshu: A communication strategy for global tourism. *Sci. Cult.* **2025**, *11*, 137–154. [\[CrossRef\]](#)
12. Vassiliadis, C.A.; Belenioti, Z.-C. Museums & cultural heritage via social media: An integrated literature review. *Tourismos* **2017**, *12*, 97–132.
13. Zollo, L.; Rialti, R.; Marrucci, A.; Ciappei, C. How do museums foster loyalty in tech-savvy visitors? The role of social media and digital experience. *Curr. Issues Tour.* **2022**, *25*, 2991–3008. [\[CrossRef\]](#)
14. Stieglitz, S.; Mirbabaie, M.; Ross, B.; Neuberger, C. Social media analytics—Challenges in topic discovery, data collection, and data preparation. *Int. J. Inf. Manag.* **2018**, *39*, 156–168. [\[CrossRef\]](#)
15. Ricciardi, R.; Manisera, M. A multilingual BERT-based classification of reviews for enhanced visitors' experience analysis. *Sci. Rep.* **2025**, *15*, 31429. [\[CrossRef\]](#)
16. Wang, Z.; Liu, W.; Sun, Z.; Zhao, H. Understanding the world heritage sites' brand diffusion and formation via social media: A mixed-method study. *Int. J. Contemp. Hosp. Manag.* **2024**, *36*, 602–631. [\[CrossRef\]](#)
17. Lan, K.; Buranaut, I.; Hsu, W.-L. Sustainable development strategies for historic building museumization based on UGC data. *J. Asian Arch. Build. Eng.* **2025**, 1–20. [\[CrossRef\]](#)
18. Zang, Z.; Fu, H.; Cheng, J.; Raza, H.; Fang, D. Digital threads of architectural heritage: Navigating tourism destination image through social media reviews and machine learning insights. *J. Asian Arch. Build. Eng.* **2025**, *24*, 4452–4469. [\[CrossRef\]](#)
19. Snyder, H. Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [\[CrossRef\]](#)
20. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Nowell, L.S.; Norris, J.M.; White, D.E.; Moules, N.J. Thematic analysis: Striving to meet the trustworthiness criteria. *Int. J. Qual. Methods* **2017**, *16*, 1–13. [\[CrossRef\]](#)
22. Agostino, D.; Brambilla, M.; Pavanetto, S.; Riva, P. The contribution of online reviews for quality evaluation of cultural tourism offers: The experience of Italian museums. *Sustainability* **2021**, *13*, 13340. [\[CrossRef\]](#)
23. Alexander, V.D.; Blank, G.; Hale, S.A. TripAdvisor reviews of London museums: A new approach to understanding visitors. *Mus. Int.* **2018**, *70*, 154–165. [\[CrossRef\]](#)
24. Balcioglu, Y.S.; Altındağ, E. An analysis of TripAdvisor reviews of top tourist attractions in Antalya, Turkey: A comprehensive analysis. *Int. J. Tour. Cities* **2026**, *12*, 534–555. [\[CrossRef\]](#)
25. Bojanic, D.; Edara, N.; Zhang, J. An examination of sentiment analysis as a tool for gathering visitor insights from online review sites for a museum. *J. Nonprofit Public Sect. Mark.* **2026**, *38*, 191–217. [\[CrossRef\]](#)

26. Burkov, I.; Gorgadze, A. From text to insights: Understanding museum consumer behavior through text mining TripAdvisor reviews. *Int. J. Tour. Cities* **2023**, *9*, 712–727. [\[CrossRef\]](#)
27. Clarizia, F.; Colace, F.; De Santo, M.; Lombardi, M.; Pascale, F. A sentiment analysis approach for evaluation of events in field of cultural heritage. In *2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS)*; IEEE: Piscataway, NJ, USA, 2018; pp. 120–127. [\[CrossRef\]](#)
28. Feng, X.; Wang, C.; Zou, T.T. Visitor experience of the grand canal national cultural park museum based on sentiment analysis algorithm. *Int. J. Electr. Electron. Eng.* **2024**, *11*, 142–150. [\[CrossRef\]](#)
29. Hariyono, H.; Wibawa, A.P.; Noviani, E.F.; Lauretta, G.C.; Citra, H.R.; Utama, A.B.P.; Dwiyanto, F.A. Exploring visitor sentiments: A study of nusantara temple reviews on Tripadvisor using machine learning. *J. Appl. Data Sci.* **2024**, *5*, 600–612. [\[CrossRef\]](#)
30. Kirilenko, A.P.; Stepchenkova, S.O.; Dai, X. Automated topic modeling of tourist reviews: Does the Anna Karenina principle apply? *Tour. Manag.* **2021**, *83*, 104241. [\[CrossRef\]](#)
31. Pinheiro, M.; Seabra, C.; Santos, N.; Caldeira, A.M.; Cravidão, F. Mapping sentiments and emotions in a UNESCO World heritage destination. *Int. J. Tour. Cities* **2026**, *12*, 610–628. [\[CrossRef\]](#)
32. Saoualih, A.; Safaa, L.; Bouhatous, A.; Bidan, M.; Perkumienė, D.; Aleinikovas, M.; Šilinskas, B.; Perkumas, A. Exploring the tourist experience of the majorelle garden using VADER-based sentiment analysis and the Latent Dirichlet Allocation Algorithm: The case of Tripadvisor Reviews. *Sustainability* **2024**, *16*, 6378. [\[CrossRef\]](#)
33. Saoualih, A.; Shen, S.; Safaa, L.; Su, Y. A thematic investigation of tourist experiential gaps in Moroccan cultural heritage museums using a sentiment-guided BERTopic text mining approach. *Tour. Recreat. Res.* **2025**, 1–25. [\[CrossRef\]](#)
34. Su, Y. Proposed museum service operations framework for addressing overcrowding: A Taipei case study. *Curator Mus. J.* **2023**, *66*, 329–350. [\[CrossRef\]](#)
35. Xu, Q.; Shih, J.-Y. Applying text mining techniques for sentiment analysis of museum visitor reviews. In *2024 IEEE 4th International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB)*; IEEE: Piscataway, NJ, USA, 2024; pp. 270–274. [\[CrossRef\]](#)
36. Colladon, A.F.; Grippa, F.; Innarella, R. Studying the association of online brand importance with museum visitors: An application of the semantic brand score. *Tour. Manag. Perspect.* **2020**, *33*, 100588. [\[CrossRef\]](#)
37. Gao, B.; Yu, S. Upgrading museum experience: Insights into offline visitor perceptions through social media trends. *Emerg. Trends Drugs Addict. Health* **2024**, *4*, 100137. [\[CrossRef\]](#)
38. Jiang, Y.; Pashkevych, K.; Bi, S. Evaluating visitor perception and spatial preferences of various museums based on machine learning from 2016 to 2024. *PLoS ONE* **2025**, *20*, e0327112. [\[CrossRef\]](#)
39. Li, Y.; Zhang, A. Innovative extraction and design application of architectural memes in ganxi former residence, Nanjing, China, based on online reviews. *Buildings* **2026**, *16*, 305. [\[CrossRef\]](#)
40. Pan, J.; Mou, N.; Liu, W. Emotion analysis of tourists based on domain ontology. In *Proceedings of the 2019 International Conference on Data Mining and Machine Learning (ICDMML 2019)*; ACM: New York, NY, USA, 2019. [\[CrossRef\]](#)
41. Peng, W.; Gao, C.; Zhu, B.; Zhu, X.; Jing, Q. Visitor behavioral preferences at cultural heritage museums: Evidence from social media data. *Buildings* **2025**, *15*, 3756. [\[CrossRef\]](#)
42. Wu, Y.; Song, X. Strategies for enhancing visitor experience with digital exhibition in museums: A case study of the Suzhou Bay Digital Art museum. In *2025 2nd International Conference on Artificial Intelligence and Future Education (AIFE 2025)*; ACM: New York, NY, USA, 2025. [\[CrossRef\]](#)
43. Xue, Y.; An, G.; Liu, W.; Kwon, K. Data-driven analysis of cultural tourism experience: Emotional journey mapping of museum visitors in China. *Asia Pac. J. Tour. Res.* **2026**, 1–25. [\[CrossRef\]](#)
44. Adriaansen, R.-J. Picturing Auschwitz. Multimodality and the attribution of historical significance on Instagram. *J. Study Educ. Dev.* **2020**, *43*, 652–681. [\[CrossRef\]](#)
45. Chang, M.; Yi, T.; Hong, S.; Lai, P.Y.; Jun, J.Y.; Lee, J.-H. Identifying museum visitors via social network analysis of instagram. *J. Comput. Cult. Herit.* **2022**, *15*, 56. [\[CrossRef\]](#)
46. Melcher, S.; Zuanni, C. Talking About the World Press Photo 20 Exhibition at the Westlicht: Analysing Communication Frames on Social Media. *Vis. Resour.* **2021**, *37*, 44–62. [\[CrossRef\]](#)
47. Rhee, B.; Pianzola, F.; Choi, G. Analyzing the museum experience through the lens of instagram posts. *Curator Mus. J.* **2021**, *64*, 529–547. [\[CrossRef\]](#)
48. Rhee, B.-A.; Pianzola, F.; Choi, J.; Hyung, W.; Hwang, J. Visual content analysis of visitors' engagement with an instagrammable exhibition. *Mus. Manag. Curatorship* **2022**, *37*, 583–597. [\[CrossRef\]](#)
49. Gerrard, D.; Sykora, M.; Jackson, T. Social media analytics in museums: Extracting expressions of inspiration. *Mus. Manag. Curatorship* **2017**, *32*, 232–250. [\[CrossRef\]](#)
50. Llewellyn, C.; Shapland, A.; Bonnet, T.; Sanderson, R.; Page, K.; Bhaugeerutty, A.; Shipp, K.; Davis, K.K.; Payne, J.; Delmas-Glass, E. Enriching exhibition scholarship. *Digit. Sch. Humanit.* **2026**, *41*, i188–i195. [\[CrossRef\]](#)

51. Chen, Y.; Zhou, W.; Cui, B.; Tong, Y. The impact of social media emotions and behavioral participation on destination image: A case of museum image mining. In *Proceedings of the 14th International Conference on Computer Science & Education (ICCSE 2019)*; IEEE: Piscataway, NJ, USA, 2019; pp. 590–595.
52. Vu, H.Q.; Luo, J.M.; Ye, B.H.; Li, G.; Law, R. Evaluating museum visitor experiences based on user-generated travel photos. *J. Travel Tour. Mark.* **2018**, *35*, 493–506. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.