Semantics-enriched spatiotemporal mapping of public risk perceptions for cultural heritage during radical events

Bai, Nan; Nourian, Pirouz; Cheng, Tao; Pereira Roders, Ana

DOI
10.1016/j.ijdrr.2024.104857

Publication date
2024

Document Version
Final published version

Published in
International Journal of Disaster Risk Reduction

Citation (APA)
Bai, N., Nourian, P., Cheng, T., & Pereira Roders, A. (2024). Semantics-enriched spatiotemporal mapping of public risk perceptions for cultural heritage during radical events. *International Journal of Disaster Risk Reduction*, *113*, Article 104857. https://doi.org/10.1016/j.ijdrr.2024.104857

International Journal of Disaster Risk Reduction 113 (2024) 104857

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

Semantics-enriched spatiotemporal mapping of public risk perceptions for cultural heritage during radical events

Nan Bai [a,b,*], Pirouz Nourian [c], Tao Cheng [d], Ana Pereira Roders [a]

[a] *UNESCO Chair in Heritage and Values: Heritage and the Reshaping of Urban Conservation for Sustainability, Department of Architectural Engineering and Technology, Delft University of Technology, 2628BL Delft, The Netherlands*

[b] *Public Administration and Policy Group, Wageningen University and Research, 6706KN Wageningen, The Netherlands*

[c] *Faculty of Geo-Information Science and Earth Observation, University of Twente, 7522NH Enschede, The Netherlands*

[d] *SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineer, University College London, Gower Street, WC1E 6BT London, United Kingdom*

ARTICLE INFO

Dataset link: 4TU Research Data Repository

Keywords:
User-generated content
Social media
Natural language processing
Topic modelling
World heritage
Heritage risk management

ABSTRACT

Cultural heritage, especially those inscribed on the UNESCO World Heritage List, is meant to be valued by mankind and protected for future generations. Triggered by radical and sometimes disastrous Heritage-Related Events (HREs), communities around the world are actively involved on social media to share their opinions and emotional attachments. This paper presents exploratory data analyses on a dataset collected from Twitter concerning HREs in World Heritage that triggered global concerns, with cases of the Notre Dame Paris fire and the Venice flood, both in 2019. The spatiotemporal patterns of tweeting behaviours of online communities before, during, and after the event demonstrate a clear distinction of activation levels caused by the HREs. The dominant emotions and topics of people during the online debate are detected and visualized with pre-trained deep-learning models and unsupervised clustering algorithms. Clear spatiotemporal dynamics can be observed from the data collected in both case studies, while each case also demonstrated its specific characteristics due to the different severity. The methodological framework proposed and the analytical outcomes obtained in this paper could be used both in urban studies to mine the public opinions about HREs and other urban events for reducing risks, and by the Geo-AI community to test spatiotemporal clustering algorithms.

1. Introduction

Triggered by radical and sometimes disastrous Heritage-Related Events (abbreviated hereinafter as HRE), communities around the world are actively involved on social media platforms, such as Twitter, Weibo, and TikTok, to share their opinions and emotional attachments [1–5]. Examples of HREs include the fire in Notre Dame de Paris, and its tower designed by Eugène Viollet-le-Duc in April 2019,[1] the terrible earthquake in Turkey and Syria destroying ancient UNESCO World Heritage (WH) sites in February 2023,[2] or the more regular occasions of floods in Venice. Management of World Heritage at disaster risks is not a new problem field, as

* Corresponding author at: UNESCO Chair in Heritage and Values: Heritage and the Reshaping of Urban Conservation for Sustainability, Department of Architectural Engineering and Technology, Delft University of Technology, 2628BL Delft, The Netherlands.

E-mail address: n.bai@tudelft.nl (N. Bai).

[1] https://www.bbc.com/news/world-europe-47971044, accessed 05 May 2023.

[2] https://www.archdaily.com/996027/a-major-earthquake-hits-turkey-and-syria-destroying-a-2000-year-old-unesco-world-heritage-site, accessed 05 May 2023.

https://doi.org/10.1016/j.ijdrr.2024.104857
Received 20 February 2024; Received in revised form 17 September 2024; Accepted 24 September 2024
Available online 28 September 2024

a detailed disaster-specific manual has already been provided by ICCROM[3] for Risk-Preparedness in 1998, following the Quebec Declaration in 1996 and the Kobe/Tokyo Declaration in 1997, both with the same topic [6,7]. In extreme cases, WH properties can be added to The List of World Heritage in Danger, calling for international attention to act against threats and to protect Outstanding Universal Value (OUV). Yet in other cases, actions still need to be taken for disaster risk preparedness, response, and recovery [6]. This line of practice has been followed by international and national organizations such as the 2007 Strategy for Reducing Risks at World Heritage Properties [8], the 2010 manual of Managing Disaster Risks for World Heritage [9], the 2011 Recommendation on the Historic Urban Landscape [10], the 2016 Guide to Risk Management of Cultural Heritage [11], and the 2024 Guidance on Post-disaster and Post-conflict Recovery and Reconstruction [12]. In the digital age, however, new challenges and opportunities are created by ubiquitous and large-scale user-generated data, which has not yet been addressed in the documents mentioned above. Social media have eased, accelerated, magnified, and even sometimes polarized the expressing and sharing mechanism [13,14]. Shortly after an event, information can spread contagiously and collective emotions can be triggered [15]. While sharing experiences, giving opinions, and expressing emotions concerning an HRE, the participating public may not be deliberately talking about the cultural significance *per se*. They, however, often reveal with others their risk perceptions about the places they have visited or live in, rich in the values and valued attributes they choose to highlight with their messages, images, and even emoji's [4,16]. Albeit bearing the risk of enhancing "mediatisation of heritage" and biasing the values [17], the dynamics of messages spreading on an intrinsic social network composed of temporally-founded heritage communities could help heritage administrators make more informed and inclusive decisions for heritage at risk [15,18]. Furthermore, all the geo-tagged and time-stamped posts on social media, as well as the corresponding HREs, are embedded in spatiotemporal and social contexts [19,20]. Aggregating information on social media and mapping the spikes on both a discrete timeline and a spatial representation could yield visualizations easily understandable for decision-makers [21]. This can help them assess the impact caused by an HRE and draw conclusions on what to do next to make responses and management plans to better reduce risks and support urban conservation [9,10,22,23].

Since an event on social media is essentially a group of semantically related posts bounded by space and time, studies have been using geo-tagged tweets to identify meaningful clusters that correspond to well-known "ground truths" and/or previously unknown real-world events [24–29]. Cheng and Wicks [24] demonstrated in London that only by using spatiotemporal information, without adding any verbal hints, meaningful events can emerge since "people tweet more than expected to describe the event and spread information". Zooming into the context of heritage, two types of HREs can conceptually exist — the events in heritage, and the heritage in events. The former are the events that particularly happen to/in a built cultural heritage, such as the Notre Dame fire and the floods in Venice mentioned at the beginning [30]. The latter are the events that happen outside of a heritage property but have a broader influence regionally or globally, eventually affecting the heritage, such as the Turkey–Syria earthquake [31] and the global COVID-19 pandemic [32–35]. Even though both can be understood as HREs, this paper only focuses on the first type (i.e., events in heritage), where reactions on social media are assumed to be about heritage itself, thus more informative for deriving values. For the second type, however, the consecutive work of Kumar [36] also demonstrated the application of crowd-sourcing and social media sensing to facilitate heritage management in disasters. Local archives have been used to study public engagement after disastrous events such as the 1966 Florence Flood [37] and 2015 Nepal Earthquake [38,39].

This paper aims to explore the online public discussions during Heritage-Related Events (HREs) and propose a methodological framework to extract useful information for heritage managers from unstructured social media data. Two case studies, i.e., the fire at Notre Dame and the flood in Venice, are demonstratively tested with the workflow (see Section 3.1). Exploratory data analyses on the spatiotemporal patterns and the semantic focuses of online discussions concerning HREs have been conducted. Three questions are to be reflected by the collected empirical data:

1. What are the spatiotemporal patterns of the posting behaviour on social media at a global scale before, during, and after a major HRE?
2. How are the emotions being expressed in social media posts and evolving over time?
3. What are the main semantic topics (as public risk perceptions) being discussed and spread and how can heritage managers learn and benefit from the outcomes for risk management concerning an HRE?

To the best of the authors' knowledge, this paper is among the first examples to provide a complete framework to study the public reactions to cultural heritage at risk and their spatiotemporal and semantic characteristics using social media data. The framework comprises pre-trained machine and deep-learning modules for the semantics, fully described with mathematical notations that are easy to generalize and update following specific needs during application. It emphasizes the cultural significance reflected in public voices during HREs and offers insights into heritage management, to prevent or mitigate disaster risks, since heritage values and cultural significance "should be the foundation on which other plans and actions are based" [9].

2. Related works

2.1. Semantic-enriched event detection on social media platforms

A closely related task to the proposed approach is the so-called "Event Detection" on social media in computer science and geographic information sciences. Clustering algorithms are used to group the posts on social media according to their proximity

[3] International Centre for the Study of the Preservation and Restoration of Cultural Property.

in diverse forms at different scales [40], in order to infer meaningful real-world events therefrom. In most such studies [24,25,27,28,40], only spatiotemporal proximity, but not the semantic proximity of posts was considered during the clustering and event detection procedure. Verbal information was only considered after spatiotemporal clustering, usually in the form of conducting an additional Topic Modelling algorithm such as Latent Dirichlet Allocation (LDA) [41] to extract the semantic features of the identified spatiotemporal clusters. Some recent studies also integrated the similarity of textual vector representations [42–44], the co-existence of relevant trending verbal entities being mentioned [45], the social distance of people [46], or other high-dimensional feature distance [47] into the clustering metrics. Moreover, in some other application scenarios other than event detection, the conventional spatiotemporal clustering algorithms on Euclidean distance can also be extended to distance on spatial or social networks [48–52].

Albeit the relevance in approaches, this manuscript does not directly involve an event detection task. Rather, it can be considered as the step both after and before an event detection. Contrary to the assumption that the location and/or time of a potential event is unknown [28,40], the occurrence of HREs are indeed prior knowledge to be studied [4]. What is interesting here is therefore not to find events from massive online discussions, but to map the spatiotemporal and semantic characteristics of known or detected radical events concerning heritage. The proposed approach can be performed after events are detected, augmenting the original brief and demonstrative case study sections [24,28,42,45] with in-depth retrospective analyses from sociological perspectives, providing insights about what was important to people during events [45]. On the other hand, additional event detection algorithms can also be performed on data collected concerning specific known HREs to detect the dynamics of sub-events. For a single event that happened in a specific heritage such as the Notre Dame fire, online discussions could span far beyond the core "heritage community" and trigger a variety of sub-topics in different places at different times. For example, people may extensively post their witness accounts of the event and share their sorrow when they first heard about the news [17,53,54]. At some specific moment, a group of people may suddenly start talking about their random guesses on who to blame for such a tragedy, which could get spread with anger as fake news [55]. In parallel, some other groups of people may start suggesting future development scenarios and proposing redesign projects, which could also get resisted and trigger another round of discussion [56,57]. All these imaginary and/or realistic scenarios could be regarded as sub-events taking a slightly different perspective of the same HRE with different spatiotemporal bounds [58–60], possibly in a hierarchical structure.

2.2. Disaster risk management for cultural heritage with social media data

Social media has been increasingly researched for natural disaster management recently, where citizens are considered as "sensors" to gather valuable information [61,62] and enable effective communications [16,63]. Wang and Ye [61] and Feng et al. [62] comprehensively reviewed the roles that social media analytics play in all phases of disaster management including preparedness, mitigation, response, and recovery. Wang and Ye [61] also summarized the combination of four core dimensions of social media data in empirical studies, i.e., space, time, content, and network structure. The dimensions can be analysed separately or simultaneously, whereas very few studies have incorporated all four dimensions [64]. As analytical frameworks for the content, supervised classification schemes are usually built under the topics of "Situation Updates" of factual information, the "Opinions/Reactions", "Sentiments/Emotions" and "Experience/Witnesses" being shared, and the "Actions" proposed for emotional, monetary, and other tangible supports [61,63]. Topic clusters discovered with unsupervised methods also discover similar schemes under different granularity [27,62]. Conceptually, the work of Kersten and Klan [27] is the most relevant existing study integrating the dimensions of space, time, and content, where spatiotemporal analyses were conducted in pairs with content mining within the area of North and South Carolina under the crisis event of a hurricane. The proposed study extends the framework to the global level with specific attention to cultural heritage, which has seldom been the focus in managing disaster risks [9].

Despite the extensive use of social media and artificial intelligence for [natural] disaster risk management, the integration in the domain of cultural heritage stays at minimum [65]. The consecutive works of Kumar [37–39] are still the only points of reference in the intersection being mentioned in literature reviews [66,67], mainly focusing on the content dimension in separate [61]. Similar to studies in natural disasters, manual coding schemes and machine learning models are used to classify the response of users into showing the "Situation" and the state of heritage after the event, conveying a "Message" with heritage as background information, recalling a past "Memory" before the event, demonstrating the "Practices" of how people used the heritage after the event [39], calling for contribution as "Action", and/or expressing the "Sentiment" for the loss of heritage [37]. More recently, an application for analysing the impact of a global event (COVID-19) on the cultural landscape through inspecting words and emojis being used by social media users can be found in [35]. Furthermore, crowd-sourcing imagery data from social media have also been used to assess and locate the physical damage of cultural heritage [68–70].

3. Data and materials

3.1. Data collection

This paper takes HREs for two UNESCO WH properties as case studies: the fire in Notre Dame de Paris, France on 15 April 2019, and the flood in Venice, Italy on 12 November 2019, both of which can be classified as "rare event" in terms of risk occurrence [11]. Whereas the severe fire in Notre Dame was described as "a catastrophe for humankind" [17,57,71], the 2019 flood in Venice was also reported as unprecedented in 50 years [69,72]. For both case studies, datasets from social media and state media have been collected, processed, and analysed in the literature [53–55,69]. They demonstrated technical innovations solving computer science problems such as spatial direction classification, geo-coding, and fake news detection. However, none of them focused on heritage.

This study collected a group of new text-based datasets for exploring the spatiotemporal patterns of public reactions during HREs. Theoretically, these existing datasets can also be integrated into the framework at a later stage as complementary information.

As a popular instant-posting social media database commonly used in event-related literature, the "`full-archive search`" endpoint of the Twitter API v2[4] was used to collect tweets about both case studies for a period of two weeks (one week before the event, and one week after), i.e., 08–22 April 2019 for the Notre Dame fire, and 05–19 November 2019 for the Venice flood. A three-step procedure was followed to collect the raw data:

1. A **local search** first queries for geo-tagged tweets in a fixed radius from the hypothetical core of HREs, i.e., 1.5 km from the Cathedral of Notre-Dame de Paris (48.852737N 2.350699E) and 8 km from the centre of Venetian Island (45.438759N 12.327145E);
2. A **global search** then queries for geo-tagged tweets that also mentioned the name of the place in several languages (not the event, thus not with words "flood" and "fire"), i.e., "Notre-Dame OR notredame OR (notre dame) OR 巴黎圣母院 OR 巴黎聖母院" and "venice OR venezia OR venedig OR venise OR venicia OR veneza OR 威尼斯", respectively. Note only a non-exhaustive subset of common languages are used during the global search for pragmatic reasons, which can be expanded in other case studies;
3. The original tweets (not necessarily geo-tagged) that are related with (being replied to, being quoted, or being the original tweet of a series of conversations) each tweet mentioned above are also collected using their IDs, as an additional round of **supplemental search**.

In all steps, collected tweet data typically include the timestamp at UTC (Coordinated Universal Time) time-zone, the pseudonymized user ID, the name-based geo-location (not always available with supplemental search), the IDs of original tweets it related with, the textual contents, as well as the language code.[5] The data collection took place respectively on 6 December 2022 (local and global searches) and 31 January - 1 February 2023 (supplemental search).

3.2. Geo-coding and pre-processing of collected data

For each case study, the tweets collected with local, global, and supplemental searches mentioned in Section 3.1 are merged to make a universal collection of data potentially related to the HREs. Different from earlier versions of Twitter APIs where latitude and longitude were provided for each geo-tagged tweet, such as in [24], the current API only provides a geo-ID indicating a name-based categorical geo-location that was selected by the user when posting. These geo-tags vary in scales ranging from micro-level POI and neighbourhood to meso-level town and city, to macro-level province and country. Since this paper aims to study tweets collected globally, a combined geo-coding (providing the latitudes and longitudes with the given name of places) and reverse geo-coding (providing the name/address of places at different administrative levels with the given latitudes and longitudes) [73] procedure to unify the resolution of geo-locations is necessary for comparison and aggregation. For pragmatic purposes, the level of "cities" is selected for such unification. Using [reverse] geo-coding Python libraries GeoPy[6] with OpenStreetMap Nominatim[7] as geocoder [74], CountryInfo,[8] and Wikipedia-API,[9] all places are merged and mapped to the city level. The more detailed geo-locations (POIs, neighbourhoods, and towns) are simply relocated in the cities where they belong, while more coarsened geo-locations (provinces, regions, and countries) are arbitrarily mapped to the capital cities (if any). As a consequence, a list of cities participating in the HRE discussions was obtained for each case study. Cities whose names were originally written in another language were translated into English using Google Translator API from the Deep Translator python library.[10] This procedure effectively provides the refined numerical geo-locations of the geo-tagged tweets.

Moreover, the raw textual tweets in different languages were fed into a pre-processing pipeline:

1. Tweet sentences are tokenized with the TweetTokenizer[11] of NLTK Python library [75];
2. The tokens are normalized by lower-casing the letters, transforming any user IDs into '`@USER`' token, changing any internet link into '`HTTPURL`' token, and de-emojizing the emojis into corresponding verbal descriptions using Emoji Python library[12];
3. The normalized tokens are joined back as sentences and translated into English using Google Translator API from the Deep Translator library.

[4] https://developer.twitter.com/en/docs/twitter-api/tweets/search/introduction, accessed 08 May 2023, Note the Twitter API rules have been updated on 29 March 2023, and the previous Academic Research Access plans were deprecated. The readers are suggested to refer to the updated version of Twitter (now X) for reproducibility.

[5] https://developer.twitter.com/en/docs/twitter-for-websites/supported-languages, accessed 08 May 2023.

[6] https://geopy.readthedocs.io/en/stable/, accessed 08 May 2023.

[7] https://github.com/osm-search/Nominatim, accessed 08 May 2023.

[8] https://github.com/porimol/countryinfo, accessed 08 May 2023.

[9] https://github.com/martin-majlis/Wikipedia-API, accessed 08 May 2023.

[10] https://deep-translator.readthedocs.io/en/latest/, Accessed 08 May 2023.

[11] https://www.nltk.org/api/nltk.tokenize.casual.html, Accessed 08 May 2023.

[12] https://carpedm20.github.io/emoji/docs/, Accessed 08 May 2023.

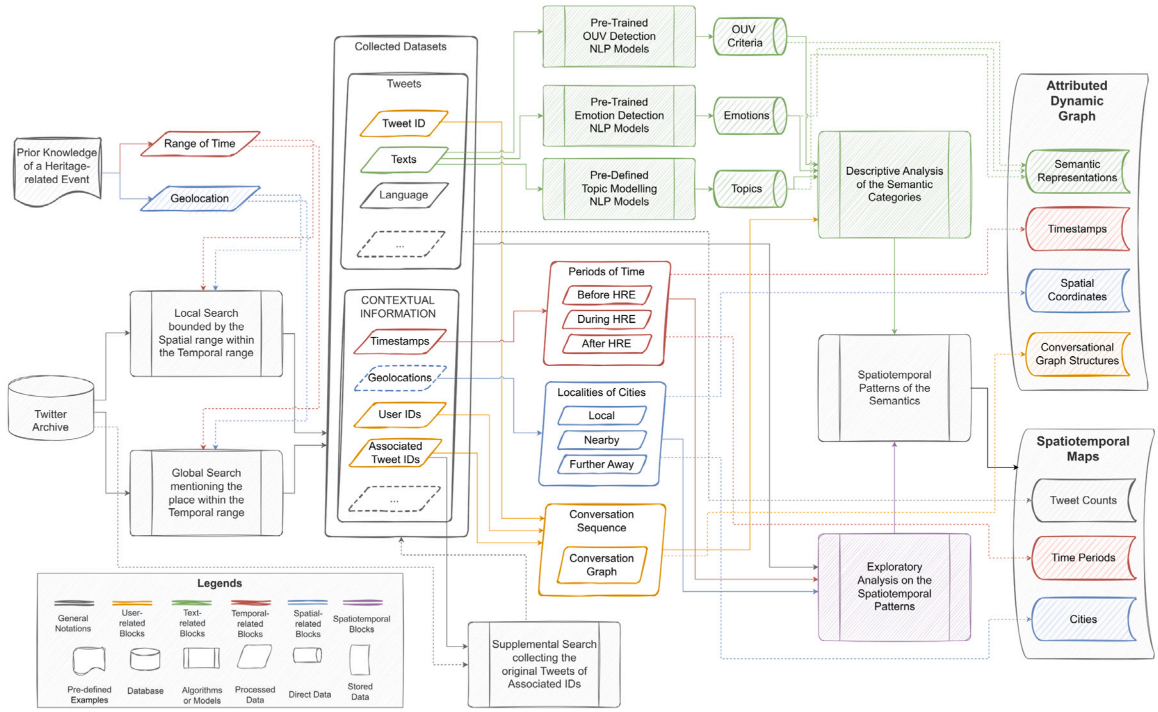

Fig. 1. The general methodological workflow proposed in this paper.

Note that no "stopwords" were removed during pre-processing, since Transformer-based Natural Language Processing (NLP) models such as BERT prefer texts to appear in their original contexts [76].

After geo-coding and pre-processing, each originally unstructured tweet can be organized as a structured tuple. Let i be the index of a generic sample of the dataset for one case study, then the tweets could be represented as a tuple $\mathfrak{d}_i := (\mathcal{S}_i, \mathcal{O}_i, \mathfrak{u}_i, \mathfrak{t}_i, \mathfrak{l}_i)$, $\mathfrak{d}_i \in \mathfrak{D} := \{\mathfrak{d}_0, \mathfrak{d}_1, \ldots, \mathfrak{d}_{K-1}\}$, where K is the number of tweets in a case study, $\mathcal{S}_i = \{s_i^{(0)}, s_i^{(1)}, \ldots, s_i^{(|S_i|-1)}\}$ is a set of normalized and translated English tweet sentences, $\mathcal{O}_i = \{\mathfrak{d}_{i'} | \mathfrak{d}_{i'} \in \mathcal{D}\}$ or $\mathcal{O}_i = \emptyset$ is the set of all collected related tweets to $\mathfrak{d}_i$, where $\mathfrak{d}_i$ referred to $\mathfrak{d}_{i'}$ in either way of interaction mentioned in Section 3.1, which can also be empty when the tweet stands alone, $\mathfrak{u}_i \in \mathcal{U}$ is a user ID that is one instance from the user set $\mathcal{U} = \{\mu_0, \mu_1, \ldots, \mu_{|\mathcal{U}|-1}\}$, $\mathfrak{t}_i \in \mathcal{T}$ is a timestamp that is one instance from the ordered set of all the unique timestamps $\mathcal{T} = \{\tau_0, \tau_1, \ldots, \tau_{|\mathcal{T}|-1}\}$ from the dataset at the level of hours, and $\mathfrak{l}_i = (\mathfrak{x}_i, \mathfrak{y}_i, \mathfrak{c}_i)$ or $\mathfrak{l}_i = \emptyset$ is the geo-location of the city where the tweet was posted, which could be empty if the tweet was not geo-tagged, including its geographical coordinate of latitude $\mathfrak{y}_i$ and longitude $\mathfrak{x}_i$, and name of the city $\mathfrak{c}_i \in C$ that is one instance from the set of unique cities $C = \{\zeta_0, \zeta_1, \ldots, \zeta_{|C|-1}\}$. For any city $\zeta_j \in C$, a corresponding geo-location (x_j, y_j) is stored as the metadata of the city. If $\mathfrak{c}_i = \zeta_j$ for a post $\mathfrak{d}_i$, it automatically entails that $\mathfrak{x}_i = x_j, \mathfrak{y}_i = y_j$.

In Notre Dame fire, the total number of tweets $K = |\mathfrak{D}| = 198,061$, the number of users $|\mathcal{U}| = 42,036$, and the number of cities $|C| = 4968$, while in Venice flood, $K = |\mathfrak{D}| = 15,641$, $|\mathcal{U}| = 3541$, and $|C| = 835$.

4. Methodology

4.1. Overview of the workflow

Fig. 1 visualizes the general framework proposed in this paper to collect, structure, and analyse the spatiotemporal dynamics of public discussion on Twitter for an HRE (Heritage-related Event). In the context of this paper, the "Prior Knowledge of a Heritage-related Event" refers to the case studies of the Notre Dame fire and Venice flood mentioned in Section 3.1, yet it can also be extended in future studies about any HRE or general social events of interest.

Through data collection, textual and contextual information of all tweets makes up a complete dataset. By inputting the textual data into several pre-trained models and pre-defined algorithms, semantic information about cultural significance, emotions, and topics is obtained as [pseudo-]labels. The contextual information of tweets is respectively used to distinguish the period relative to the HRE (before, during, and after), determine the position of cities relative to the city where the HRE happened (same city as local, nearby cities within a given radius, and global cities further away), and construct a directed network marking the conversation sequence on Twitter. After that, the tweets in all cities along the timeline are analysed through exploratory data analysis to describe the general spatiotemporal pattern and behavioural changes on Twitter during HRE. Furthermore, descriptive analyses are conducted on the semantic labels to uncover the dynamic associations among the entailed cultural significance, expressed emotions, and discussed key topics with the spatial and temporal bounds. In summary, the proposed workflow covers all four dimensions of space, time, content, and (to some extend) network in the mapping of public reactions and risk perceptions during HREs, as mentioned in Section 2.2, both in separation and in combination.

4.2. Dynamics in contextual information

The tweet counts are first aggregated. Temporally, the tweets are counted every hour as in $\mathcal{T}$ from one week before the event till one week after, resulting a vector $\boldsymbol{t} := [t_k]_{|\mathcal{T}|\times 1} \in \mathbb{N}^{|\mathcal{T}|\times 1}, t_k = |\{\mathfrak{d}_i | \mathfrak{t}_i = \tau_k\}|$ to draw a timeline on the volume and intensity of the discussion globally. Since some original posts (collected through the supplemental search) spanned far before the HRE, those posted before a certain date (08 April 2023 and 05 November 2023, respectively) were filtered out from further analysis. Spatially, the tweets are counted in the level of cities as in $\mathcal{C}$, resulting a vector $\boldsymbol{c} := [c_j]_{|\mathcal{C}|\times 1} \in \mathbb{N}^{|\mathcal{C}|\times 1}, c_j = |\{\mathfrak{d}_i | \mathfrak{c}_i = \zeta_j\}|$, and further grouped to the level of countries. Spatial and temporal intervals are then combined to further separate the posts. Specifically, the set of timestamps $\mathcal{T}$ is divided into $\mathcal{T}_\mathrm{B} \subset \mathcal{T}$ before the main HREs, $\mathcal{T}_\mathrm{D} \subset \mathcal{T}$ during the HREs up to four days after the event, and $\mathcal{T}_\mathrm{A} \subset \mathcal{T}$ after the main HREs upon one week after. And the set of cities $\mathcal{C}$ is divided into $\mathcal{C}_0 = \{\zeta_0\} \subset \mathcal{C}$ which is the "host" city of the HRE (Paris or Venice), $\mathcal{C}_1 \subset \mathcal{C}$ which contains cities from the same country (France or Italy), and $\mathcal{C}_2 \subset \mathcal{C}$ containing cities from elsewhere. This categorization of time is referred to as "Periods" and that of cities as "Locality" in this paper.

Considering the different Periods, the vector $\boldsymbol{c}$ can be disaggregated into three vectors $\boldsymbol{c}_\mathrm{B}, \boldsymbol{c}_\mathrm{D}, \boldsymbol{c}_\mathrm{A} \in \mathbb{N}^{|\mathcal{C}|\times 1}$, where $\boldsymbol{c}_\mathrm{B} + \boldsymbol{c}_\mathrm{D} + \boldsymbol{c}_\mathrm{A} = \boldsymbol{c}$, respectively counting the number of tweets posted in each city before, during, and after HREs, the entries of which can be 0 when no tweets are posted in a city for specific periods, very common in $\mathcal{C}_2$ cities all over the world before the HRE. The entries of the vectors $\boldsymbol{c}_\mathrm{B}, \boldsymbol{c}_\mathrm{D}, \boldsymbol{c}_\mathrm{A}$ are sorted in descending order, generating ranks of the cities in each period, where the ranks of cities with a tie are arbitrarily assigned. Then the numbers of tweets in all cities in each period are plotted against their rankings $\boldsymbol{n} = [1, 2, 3..., |\mathcal{C}|]^\mathrm{T}$ in a log–log scale, resulting in rank–size plots. This is to check if the fat-tailed distribution in the seminal Zipf's Law or the more general power law still holds in terms of tweeting behaviour in HREs and if there is a pattern shift among the different periods [77,78]. Moreover, for each tweet $\mathfrak{d}_i$, the geodesic distance (the arc length on Earth surface) $\boldsymbol{d} := [d_i]_{K\times 1} \in \mathbb{R}^{K\times 1}$ of the city $\mathfrak{c}_i$ where it is posted and the city ζ_0 where the HRE actually happened can be computed, using their respective coordinates $(\mathfrak{x}_i, \mathfrak{y}_i)$ and (x_0, y_0). The distributions of vector $\boldsymbol{d}$ are also plotted while distinguishing the tweets in different periods (thus 3 distributions for each case study). The geodesic distance is computed using GeoPy Python library.

Conversation sequence of Tweets is modelled as a directed multi-graph $\mathcal{G}^{\mathrm{MULT}} = (\mathcal{V}, \{\mathcal{E}^{\mathrm{CONV}}, \mathcal{E}^{\mathrm{USER}}\})$. Two types of links are considered:

1. For any tweet $\mathfrak{d}_i$, as long as $\mathcal{O}_i$ set of related tweets is not empty, the tweet $\mathfrak{d}_i$ and all $\mathfrak{d}_{i'} \in \mathcal{O}_i$ are added to the node set $\mathcal{V}$, and the links pointing from the tweet to its associated ones $(\mathfrak{d}_i, \mathfrak{d}_{i'})$ are added to the link set $\mathcal{E}^{\mathrm{CONV}}$.
2. For any user $\mu_u \in \mathcal{U}$, all the tweets posted by them are assembled as $\{\mathfrak{d}_i | \mathfrak{u}_i = \mu_u\}$ and added to the node set $\mathcal{V}$, which are then ranked in chronological order. The links pointing from the later tweets to their immediate temporal neighbour are added to the link set $\mathcal{E}^{\mathrm{USER}}$.

The nodes inherit all the data from the tweets as node attributes, including text, language, user information, timestamp, period, and [possibly] latitude, longitude, and locality of its posting city. In the end, the node sets $\mathcal{V} \subseteq \mathcal{D}$ became the bases for further semantic analyses, which were smaller than the raw tweet sets $\mathcal{D}$. In Notre-Dame fire, the number of tweets appearing as nodes in the conversational network $|\mathcal{V}| = 179,758$ (79.3% of $|\mathcal{D}|$), the number of conversational links including interactions of different users and consecutive posting of the same users $|\mathcal{E}^{\mathrm{CONV}} \bigcup \mathcal{E}^{\mathrm{USER}}| = 221,285$, while in Venice flood, $|\mathcal{V}| = 11{,}961$ (76.5% of $|\mathcal{D}|$), and $|\mathcal{E}^{\mathrm{CONV}} \bigcup \mathcal{E}^{\mathrm{USER}}| = 12{,}106$.

4.3. Semantics of tweets

For each tweet $\mathfrak{d}_i \in \mathcal{V} \subseteq \mathcal{D}$, the translated English sentences $\mathcal{S}_i$ are fed into pre-trained models and pre-defined algorithms to obtain their semantic meanings.

4.3.1. Entailed cultural significance

Outstanding Universal Value (OUV) is the core concept defining the cultural significance of UNESCO WH [79,80]. Ten selection criteria (six cultural and four natural) were also defined to justify why the cultural significance of a WH property is so exceptional, that it transcends national boundaries and is to be preserved for future generations of all humanity. The justification of the OUV of a WH property (i.e., the Statement of OUV) during/after inscription highlights the main heritage values and attributes to take as reference during disasters and while assessing the risks [9,81]. Conceptual frameworks on cultural significance, theorizing values and attributes, are not bounded to specific governance levels (e.g., local – national - international), heritage categories (e.g., intangible – tangible), and legal status (e.g., listed - unlisted) [21,82–84]. They can be used in various settings, such as to identify cultural significance, and to assess the impact and risks for development projects and/or disasters. For demonstrative purposes, however, the BERT [76] and ULMFiT [85] models pre-trained and finetuned with the UNESCO Statements of OUV in [86], which are the only available open-source machine learning model to analyse cultural significance so far, are used to predict the most probable OUV selection criteria mentioned in the tweets, such that:

$$\boldsymbol{y}_i^{\mathrm{BERT}} = g_{\mathrm{BERT}}(\mathcal{S}_i | \boldsymbol{\Theta}_{\mathrm{BERT}}), \boldsymbol{y}_i^{\mathrm{ULMFiT}} = g_{\mathrm{ULMFiT}}(\mathcal{S}_i | \boldsymbol{\Theta}_{\mathrm{ULMFiT}}),$$

$$\boldsymbol{Y}^{\mathrm{OUV}} := [\boldsymbol{y}_i^{\mathrm{OUV}}]_{11\times|\mathcal{V}|}, \boldsymbol{y}_i^{\mathrm{OUV}} = \frac{\boldsymbol{y}_i^{\mathrm{BERT}} + \boldsymbol{y}_i^{\mathrm{ULMFiT}}}{2}. \tag{1}$$

where $\boldsymbol{g}_{(*)}$ is the end-to-end function of pre-trained models, $\boldsymbol{\Theta}_{(*)}$ is the parameter of the models, and $\boldsymbol{y}_i^{(*)} \in [0,1]^{11\times 1}$ is an 11-dimensional logit vector as soft-label predictions. Let top-n$(\boldsymbol{l}, n)$ denote the function returning the index set of the largest n elements of a vector $\boldsymbol{l}$, let max$(\boldsymbol{l}, n)$ denotes the function returning the value of the n_{th} largest element of vector $\boldsymbol{l}$, and let IoU$(\mathcal{A}, \mathcal{B}) := |\mathcal{A} \bigcap \mathcal{B}|/|\mathcal{A} \bigcup \mathcal{B}|$ denotes the Intersection over Union (Jaccard Index) of any two generic sets $\mathcal{A}, \mathcal{B}$, then the confidence and [dis-]agreement of both models for top-n predictions could be computed as:

$$\begin{aligned}
\boldsymbol{K}^{\mathrm{OUV}} &:= [\boldsymbol{\kappa}_i^{\mathrm{OUV}}]_{2\times|\mathcal{V}|}, \boldsymbol{\kappa}_i^{\mathrm{OUV}} := [\kappa_i^{\mathrm{OUV}(0)}, \kappa_i^{\mathrm{OUV}(1)}]^{\mathrm{T}}, \\
\kappa_i^{\mathrm{OUV}(0)} &= \sum_{n_0=1}^{n} \frac{\max(\boldsymbol{y}_i^{\mathrm{BERT}}, n_0) + \max(\boldsymbol{y}_i^{\mathrm{ULMFiT}}, n_0)}{2}, \\
\kappa_i^{\mathrm{OUV}(1)} &= \mathrm{IoU}\left(\text{top-n}(\boldsymbol{y}_i^{\mathrm{BERT}}, n), \text{top-n}(\boldsymbol{y}_i^{\mathrm{ULMFiT}}, n)\right),
\end{aligned} \tag{2}$$

where $\kappa_i^{\mathrm{OUV}(0)}$ denotes the average accumulated top-n confidence (probability) of both models, and $\kappa_i^{\mathrm{OUV}(1)}$ denotes the agreement of the models in their categorical predictions. Since it was noted that the models work better with top-3 predictions [86], only the tweets that are predicted by both models with higher top-3 confidence of .75 and top-3 agreement of .50 are considered as expressing information related to cultural significance, making up a subset of tweet nodes $\mathcal{V}^{\mathrm{OUV}} \subset \mathcal{V} \subset \mathcal{D}$. Note this approach is similar to Bai et al. [20], where the filtering thresholds of confidence and agreement are directly inherited. Changing the thresholds would further alter the proportion and quality of posts to be kept. After filtering, the number of tweets classified as mentioning cultural significance in Notre-Dame fire is $|\mathcal{V}^{\mathrm{OUV}}| = 61{,}550$ (34.2% of $|\mathcal{V}|$), and for Venice flood $|\mathcal{V}^{\mathrm{OUV}}| = 3628$ (30.3% of $|\mathcal{V}|$). And the predicted categorical top-3 [pseudo-]labels of OUV selection criteria for the tweets can be described as an array of sets:

$$\mathcal{Y}^{\mathrm{OUV}} = \left[\mathcal{Y}_i^{\mathrm{OUV}}\right] = \left[\left\{\text{top-n}(\boldsymbol{y}_i^{\mathrm{OUV}}, 3) | \mathfrak{d}_i \in \mathcal{V}^{\mathrm{OUV}}\right\} \text{ or } \varnothing\right]. \tag{3}$$

4.3.2. Expressed emotions

Sentiments and emotions are important semantic aspects to research from social media during disaster crises [61,62]. Pre-trained models on sentiment analysis and emotion detection tasks are conventionally used [87–90]. Both the NLP tasks have been applied in the analysis of User-Generated Content for heritage and tourism studies [91–96]. Whereas sentiment analysis only classifies texts into a polarity of negative, positive, and neutral, emotion detection considers the full spectral of 6 basic human emotions by Paul Ekman [97], i.e., joy (happiness), sadness, fear, disgust, anger, and surprise. Preliminary studies show that some Emotion Detection models are not robust enough with different input data to produce consistent predictions, even if outputting high confidence. Therefore, this study integrates predictions by several models, and only keeps the ones with high consistency across tasks, similar to ensemble learning [98]. On the one hand, pysentimiento Python toolkit is used to predict both the sentiment[13] and the emotion[14] categories of the tweets [99–101]. On the other hand, additional models for sentiment analysis[15] and emotion detection[16] are respectively used, both finetuned with BERTweet as base models [100,102]. All predictions are conducted with the "text classification" pipeline in Huggingface Transformer Python library [103]. It is worth noting that pysentimiento enables an additional class of "others" aside from the 6 basic emotions, not forcing the model to predict one emotion category if the sentence is considered as neutral.

Similar to Eq. (2), the emotion logic vectors $\boldsymbol{y}_i^{\mathrm{EM}(0)} \in [0,1]^{7\times 1}, \boldsymbol{y}_i^{\mathrm{EM}(1)} \in [0,1]^{6\times 1}$ and sentiment vectors $\boldsymbol{y}_i^{\mathrm{SE}(0)}, \boldsymbol{y}_i^{\mathrm{SE}(1)} \in [0,1]^{3\times 1}$ computed by respective models could be accompanied with measures to ensure consistency:

$$\begin{aligned}
\kappa_i^{\mathrm{EM}} &= \begin{cases} \text{top-n}(\boldsymbol{y}_i^{\mathrm{EM}(1)}, 1) = \text{top-n}(\boldsymbol{y}_i^{\mathrm{EM}(0)}, 1) & \text{if top-n}(\boldsymbol{y}_i^{\mathrm{EM}(0)}, 1) \neq \text{‘others’} \\ \text{top-n}(\boldsymbol{y}_i^{\mathrm{EM}(1)}, 1) \in \text{top-n}(\boldsymbol{y}_i^{\mathrm{EM}(0)}, 2) & \text{otherwise,} \end{cases} \\
\kappa_i^{\mathrm{SE}} &= \begin{cases} 1 & \text{if top-n}(\boldsymbol{y}_i^{\mathrm{SE}(0)}, 1) = \text{‘NEU’ or top-n}(\boldsymbol{y}_i^{\mathrm{SE}(1)}, 1) = \text{‘NEU’} \\ \text{top-n}(\boldsymbol{y}_i^{\mathrm{SE}(1)}, 1) = \text{top-n}(\boldsymbol{y}_i^{\mathrm{SE}(0)}, 1) & \text{otherwise,} \end{cases} \\
\boldsymbol{k}^{\mathrm{EMS}} &= [\kappa_i^{\mathrm{EMS}}]_{1\times|\mathcal{V}|}, \kappa_i^{\mathrm{EMS}} = \kappa_i^{\mathrm{EM}} \wedge \kappa_i^{\mathrm{SE}}, \text{where } \kappa_i^{\mathrm{EMS}}, \kappa_i^{\mathrm{EM}}, \kappa_i^{\mathrm{SE}} \in \{0, 1\}.
\end{aligned} \tag{4}$$

The emotion labels are only considered as consistent ($\kappa_i^{\mathrm{EM}} = 1$) when the top-1 predictions of both models are the same, or in case pysentimiento considers a tweet as neutral, the second most significant emotion is the same as the other model. And the sentiment labels are considered as similar ($\kappa_i^{\mathrm{SE}} = 1$) when the top-1 predictions of both models are the same or when either model predicts "NEU" (neutral) polarity. Only when the tweet has both a consistent emotion detection and a similar sentiment detection result, the emotion prediction is considered valid ($\kappa_i^{\mathrm{EMS}} = 1$), resulting in a subset of tweets $\mathcal{V}^{\mathrm{EMS}} \subset \mathcal{V} \subset \mathcal{D}$ expressing emotions. After filtering with consistency, the tweets classified as expressing consistent emotions in Notre-Dame fire are $|\mathcal{V}^{\mathrm{EMS}}| = 93{,}616$ (52.1% of $|\mathcal{V}|$), among which 27,375 displayed an explicit emotion other than 'others' (15.2% of $|\mathcal{V}|$), while in Venice flood, the numbers are

[13] https://huggingface.co/pysentimiento/robertuito-sentiment-analysis, accessed 12 May 2023.
[14] https://huggingface.co/pysentimiento/robertuito-emotion-analysis, accessed 12 May 2023.
[15] https://huggingface.co/finiteautomata/bertweet-base-sentiment-analysis, accessed 12 May 2023.
[16] https://huggingface.co/Emanuel/bertweet-emotion-base, accessed 12 May 2023.

respectively $|\mathcal{V}^{\text{EMS}}| = 6235$ (also 52.1% of $|\mathcal{V}|$) and 1573 with explicit emotions (13.2% of $|\mathcal{V}|$). The predicted categorical emotion and sentiment [pseudo-]labels of each tweet can therefore be described as an array of sets:

$$\mathcal{Y}^{\text{EMS}} = \left[\mathcal{Y}_i^{\text{EMS}}\right] = \left[\left\{\text{top-n}(\boldsymbol{y}_i^{\text{EM}(1)}, 1), \text{top-n}(\boldsymbol{y}_i^{\text{SE}(1)}, 1) | \kappa_i^{\text{EMS}} = 1 \wedge \mathfrak{d}_i \in \mathcal{V}^{\text{EMS}}\right\} \text{ or } \emptyset\right]. \tag{5}$$

4.3.3. Discussed topics

In addition to supervised semantic categories such as cultural significance, semantics, and emotions, unsupervised topic models have been broadly used to identify semantic structures previously unknown and undefined for disaster management [27,61]. BERTopic[17] Python library is used to conduct unsupervised topic modelling [104]. BERTopic is a modular pipeline with components that is context-aware and uses state-of-the-art sentence embeddings of pre-trained models, which differs the conventional topic modelling algorithm LDA [41] that only considers documents as bag-of-words. For each HRE, the BERTopic pipeline is called upon the translated English tweets $\mathcal{S}_i$ to generate topics. Each topic needs to appear more than 45 times for Notre-Dame fire and 25 times for Venice flood. The number of topics was not determined arbitrarily but was merged automatically with HDBSCAN [105]. Afterwards, all the generated topics (denoted as $\mathcal{Z} = \{z_m | m = -1, 0, 1, \dots, |\mathcal{Z}| - 2\}$) are checked manually by an expert to first qualitatively ensure the validity of obtained topics and then select a subset $\mathcal{Z}^s$ presumably informative for heritage management, such as incidence description, emotion expressions, and call for actions. The implementation details of the topic modelling procedure using BERTopic can be found in Appendix A. Other than the mentioned parameters, default configuration is generally implemented. Note the clustering quality was only controlled via the internal components of BERTopic such as HDBSCAN and c-TF–IDF [104]. No explicit hyper-parameter tuning was performed as finding the optimal clustering algorithm is out of the current scope. The outcome of BERTopic is effectively a probability distribution for each tweet over all topics (including the 'noise' topic z_{-1}):

$$\boldsymbol{y}_i^{\text{TOP}} = \left[y_{i,m}^{\text{TOP}}\right]_{|\mathcal{Z}|\times 1} \in [0,1]^{|\mathcal{Z}|\times 1}, \mathbf{1}_{|\mathcal{Z}|\times 1}^{\text{T}} \boldsymbol{y}_i^{\text{TOP}} = \sum y_i^{\text{TOP}} = 1, \tag{6}$$

where $y_{i,m}^{\text{TOP}}$ refers to the probability of the i_{th} tweet being categorized as the m_{th} topic within $\mathcal{Z}$, and $\mathbf{1}$ is a vector of all 1s. Keeping only the predictions with high confidence ($y_{i,m}^{\text{TOP}} > 0.5$ as an arbitrary cutting point) for the interesting topics $z_m \in \mathcal{Z}^s$, a subset of tweets $\mathcal{V}^{\text{TOP}} \subset \mathcal{V} \subset \mathcal{D}$ referring to heritage-informative topics can be obtained. For Notre-Dame fire, the number of obtained topics after topic modelling is $|\mathcal{Z}| = 260$, the number of informative key topics $|\mathcal{Z}^s| = 57$ (52.1% of $|\mathcal{Z}|$), and the number of tweets referring to key topics $|\mathcal{V}^{\text{TOP}}| = 77{,}007$ (42.8% of $|\mathcal{V}|$), among which 8206 are not within the 'noise' topic z_{-1} (4.6% of $|\mathcal{V}|$). And for Venice flood, the numbers are respectively $|\mathcal{Z}| = 45$, $|\mathcal{Z}^s| = 22$ (48.9% of $|\mathcal{Z}|$), $|\mathcal{V}^{\text{TOP}}| = 5515$ (46.1% of $|\mathcal{V}|$), among which 1836 are not within the 'noise' topic (15.3% of $|\mathcal{V}|$). The eventual topic [pseudo-]labels of each tweet can be described as an array of sets:

$$\mathcal{Y}^{\text{TOP}} = \left[\mathcal{Y}_i^{\text{TOP}}\right] = \left[\left\{z_m | y_{i,m}^{\text{TOP}} > 0.5 \wedge \mathfrak{d}_i \in \mathcal{V}^{\text{TOP}} \wedge z_m \in \mathcal{Z}^s\right\} \text{ or } \emptyset\right]. \tag{7}$$

After obtaining all semantic [pseudo-]labels $\mathcal{Y}^{\text{OUV}}, \mathcal{Y}^{\text{EMS}}, \mathcal{Y}^{\text{TOP}}$, the timelines demonstrating the temporal development of each type of semantics are visualized, for which the tweets under different Periods and Localities are also counted. Both steps are similar to the procedure that has been mentioned in Section 4.2.

5. Results

5.1. Spatiotemporal patterns of tweeting behaviour

The temporal distribution of tweets (vector $\boldsymbol{t}$) is visualized in Fig. 2. It shows that Notre-Dame is generally tweeted more than Venice, while HREs triggered the discussion and raised the scale of tweeting behaviour to a significantly higher level. However, the peaks also vanished quickly after a few days, dropping to the scale before the event. This effect is more obvious in Notre-Dame (almost 10 folds) than in Venice (about 3 folds), possibly because even though exceptionally severe in 50 years, Venice undergoes and recovers from floods almost annually, making this HRE incomparable with the fire in Notre-Dame that shocked the entire world drastically.

Both the aggregated general spatial patterns (vector $\boldsymbol{c}$) and the disaggregated ones with different periods (vectors $\boldsymbol{c}_{\text{B}}, \boldsymbol{c}_{\text{D}}, \boldsymbol{c}_{\text{A}}$) are visualized in Figs. 3 and 4. Except that the Notre-Dame fire had a much larger scale than Venice flood, spreading to more cities worldwide, the both cases demonstrated similar patterns. The figures confirmed the hypothesis in [4] that the online discussions on Twitter triggered by HREs would go beyond the geographical boundaries, forming a global community caring about WH. Naturally, the tweets posted from the same country (France or Italy) made the largest contribution to the discussion, composing almost half of the tweet-scape, while countries nearby (e.g., EU countries) and far away also participated substantially. Interestingly, United States, United Kingdom, France, and Italy all entered the top-5 in both cases, indicating the concentration of heritage-aware people. However, this spatial pattern also probably correlates with the number of active Twitter users, the main target group, and the major purpose of usage in different regions. For example, the voice of China is significantly missing, since people there mainly use Weibo and WeChat for instant reaction and personal blogging. Moreover, through data exploration, the large amount of "Venice" discussed in United States before and after the flood turned out to be related to somewhere in Florida with the same name (Venice Beach), thus an unexpected outlier not excluded during data collection.

[17] https://maartengr.github.io/BERTopic/index.html, accessed May 13 2023.

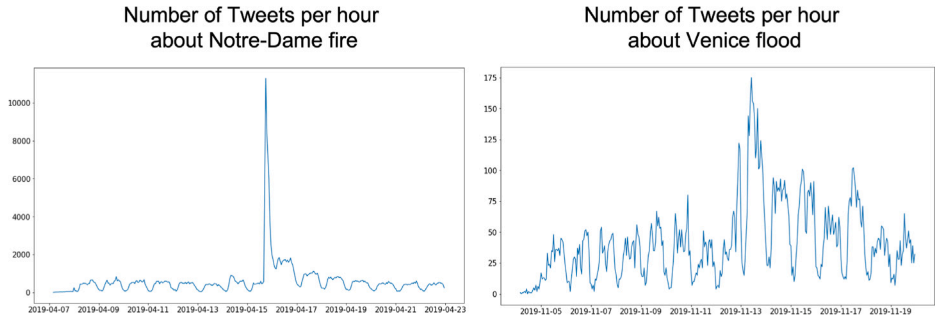

Fig. 2. The temporal pattern of tweets per hour throughout the data collection period concerning heritage-related events.

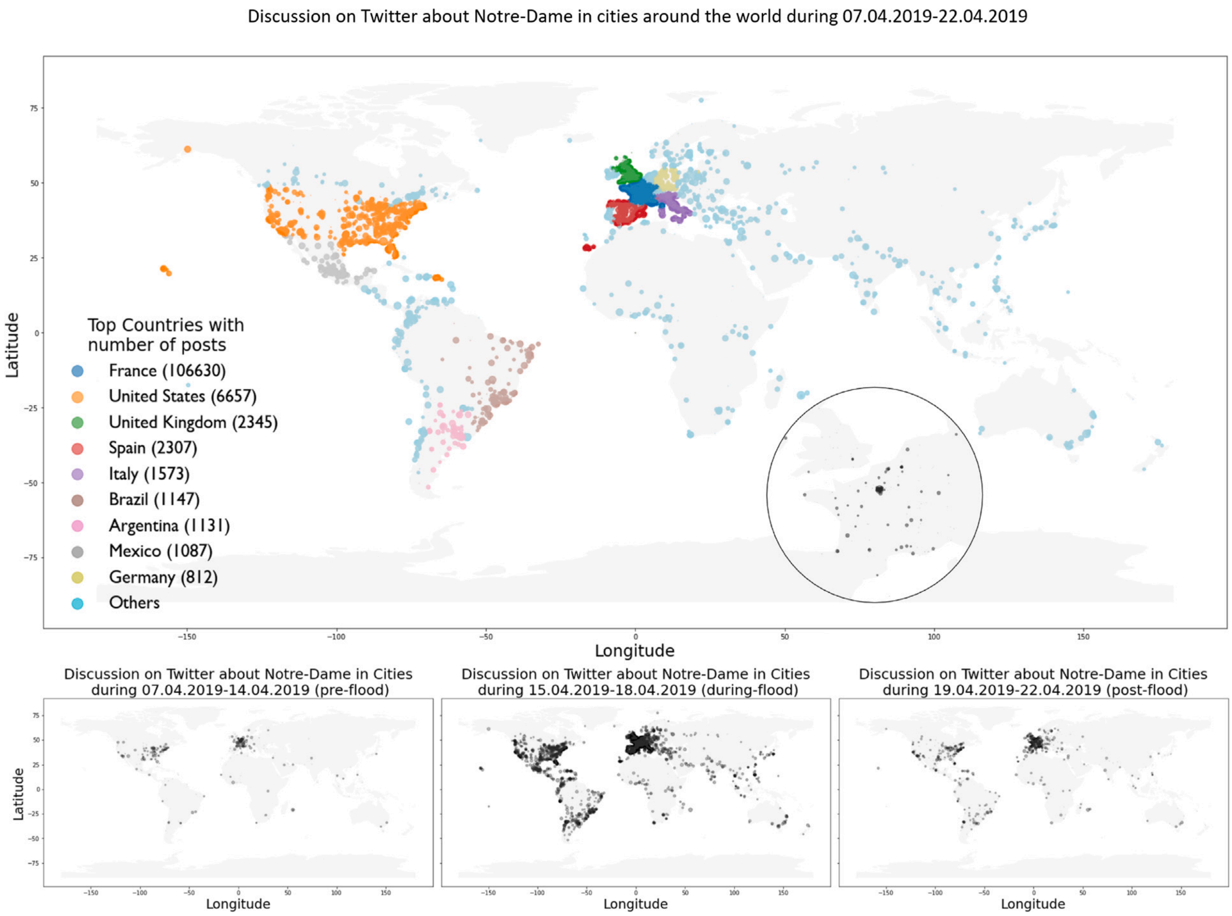

Fig. 3. The global spatial pattern of tweets throughout the data collection period in Notre-Dame fire. The larger the size of a node, the more tweets are located in the city it represents. Nodes are coloured by the top 9 countries contributing to the tweet-scape. The spatial pattern is further disaggregated in periods before, during, and after the events. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The spatiotemporal patterns of cities tweeting before, during, and after HREs could be compared by plotting the ranked vectors $\boldsymbol{c}_{\mathrm{B}}, \boldsymbol{c}_{\mathrm{D}}, \boldsymbol{c}_{\mathrm{A}}$ against their ranking $\boldsymbol{n}$ in a log–log scale, as shown in Fig. 5. The ubiquitous quasi-linear pattern of the rank–size plots in social sciences and urban studies indicating a power law of the sizes can again be observed, except for the extremely huge number of tweets (heavy head) in the highest-ranking city (i.e., the city where the event happened). By excluding the highest-ranking city, more reasonable lines can be fitted. The online participation spread to more cities globally during HREs with the longest tails, while the posting behaviour of the post-event period did not yet fully recover to the pre-event level. Even though the tweets almost always counted highest during HREs in a city, followed by after and then before HREs, this is not the case for the highest-ranking cities.

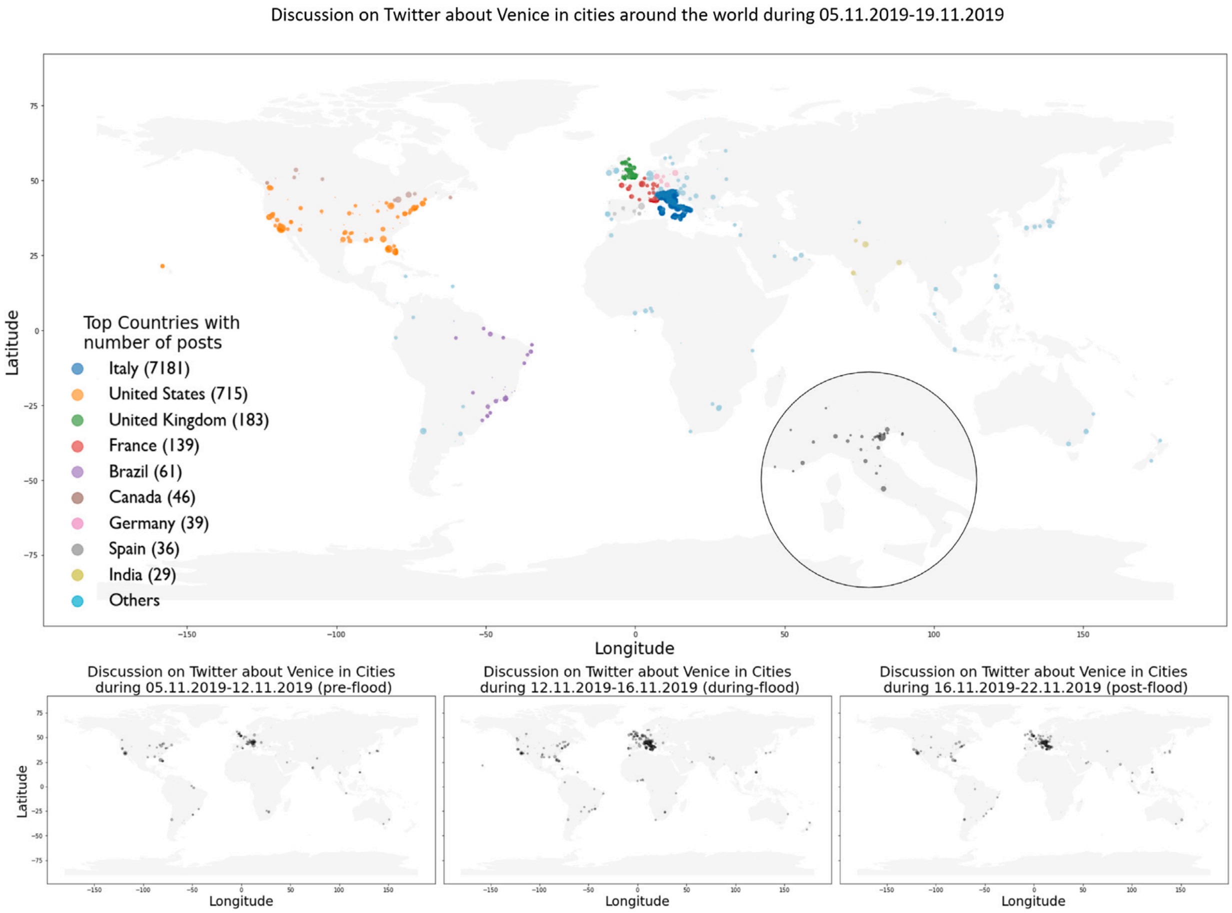

Fig. 4. The global spatial pattern of tweets throughout the data collection period in Venice flood.

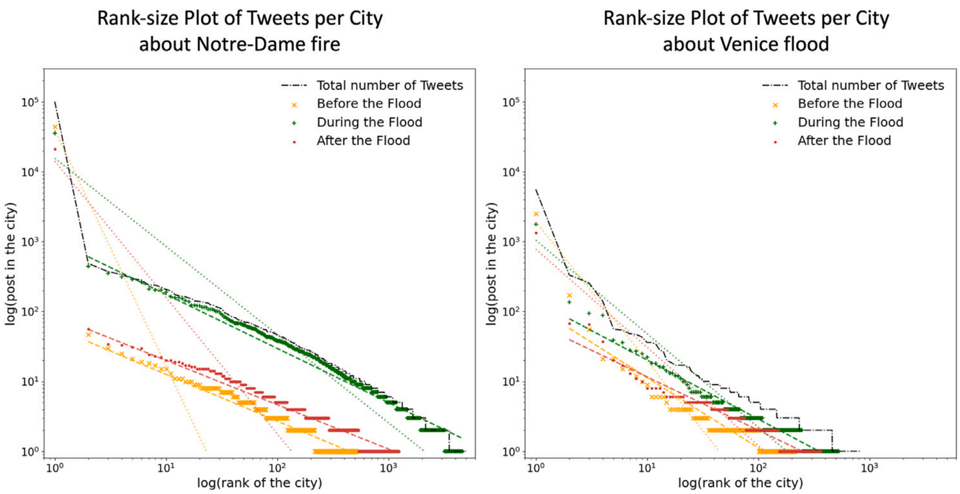

Fig. 5. The log-scale rank–size plot of tweets per city in periods before, during and after the events in Notre-Dame and Venice. Two lines are fitted to the points using the Maximum Likelihood algorithm, while the dotted ones include the highest-ranking city for the fitting, and the dashed lines exclude them.

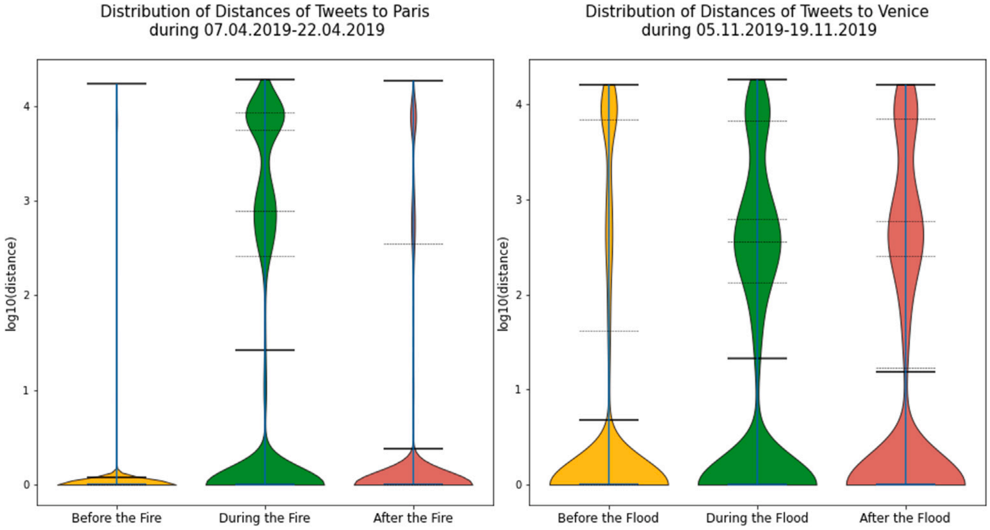

Fig. 6. The violin plots showing the distribution of distances of tweets to the core of an HRE before, during and after the event in Notre-Dame and Venice. The mean, 60%, 70%, 80%, and 90% deciles are visualized.

Table 1
Post-hoc Mann–Whitney U-tests comparing the median of ordinal variable Locality before, during, and after HREs.

Case study		Notre-Dame fire			Venice flood		
Statistics		U-value	p-value	RBC[a]	U-value	p-value	RBC[a]
Before HREs	During HREs	840,735,944.0	<.001	−.409	4,090,061.5	<.001	−.230
Before HREs	After HREs	602,097,762.0	<.001	−.097	4,135,609.0	<.001	−.173
During HREs	After HREs	511,615,038.5	<.001	.318	3,532,576.5	<.001	.053

[a] Rank Biserial Correlation (the difference between the proportions of favourable and unfavourable evidence) as effect size.

For them, the posts before HREs are even higher than that during HREs in both case studies, probably due to the number of days in each period (7 days before, 4 days during, and 4 days after HREs).

Furthermore, the distribution of the distances $\boldsymbol{d}$ from each tweet node (if $\mathfrak{c}_i \neq \emptyset$) to the city where the HREs happened is visualized in Fig. 6, demonstrating the changes of global engagement. By testing on the ordinal variable of locality of the tweets ('0' posted in the same city, '1' from the same country, or '2' from far beyond), Kruskal–Wallis H-tests showed a significant difference across the periods (before, during, and after HREs), $H(2) = 26,449.3, p < .001$ in Notre-Dame and $H(2) = 374.2, p < .001$ in Venice. Post-hoc two-tailed Mann–Whitney U-tests in Table 1 showed significant differences in the medians of locality among all pairs of HREs periods, where the period during HREs significantly and consistently showed the broadest span of locality. Almost all comparisons are justified with a medium effect size with an absolute value equal to or larger than 0.1, except for the difference between during and after HREs in Venice flood. Both statistics are calculated using Pingouin[18] Python library.

5.2. Detected cultural significance, emotions, and key topics

As mentioned already in Section 4.3, the tweets with non-empty pseudo-labels for OUV selection criteria $\mathcal{Y}^{\mathrm{OUV}}$, emotions $\mathcal{Y}^{\mathrm{EMS}}$, and key topics $\mathcal{Y}^{\mathrm{TOP}}$ all count less than the entire dataset. The relations of overlapping (in terms of the number and proportion of tweets) of the three types of semantic labels are visualized in Fig. 7. The proportions of the two case studies are similar with a significant Spearman correlation of $\rho = .976, p < .001$, where pure emotional expressions without mentioning cultural significance or key topics are consistently the majority, and tweets with all three labels are always the minority. More tweets had overlapping labels than standing alone, implying the associative nature of the cultural significance, emotions, and topics, classifying the tweets from a different perspective.

Among the tweets with $\mathcal{Y}^{\mathrm{OUV}}$ possibly mentioning cultural significance, Criterion (vi) about people's association and activity, Criterion (iii) about the testimony of a [religious/cultural] tradition, and Criterion (iv) about the architectural typology are consistently the three most significant ones, both in Notre-Dame and in Venice. Then there always followed Criterion (i) about a masterpiece, Criterion (ii) about values and influence, and Criterion (vii) about natural beauty, in a slightly different order. Even though "Paris, Banks of the Seine" as a WH property including Notre-Dame was only officially justified with OUV selection criteria

[18] https://pingouin-stats.org/build/html/index.html, accessed 16 May 2023.

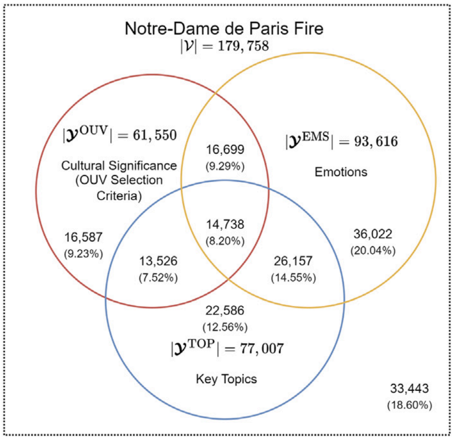

Fig. 7. The Venn Diagram of the number of tweets with each type of semantic label.

(i)(ii)(iv), and "Venice and Its Lagoon" was only justified with (i)(ii)(iii)(iv)(v)(vi), the tweets related to unjustified criteria, such as criterion (vii) in both cases, are not surprising. This is because the NLP models employed in this study read the sentences literally and return the best-matching OUV selection criteria with the knowledge of the wordings in the whole WH List. The tweets detected as relevant with $\mathcal{Y}^{\mathrm{OUV}}$ followed a logical order of how laypeople perceive the cultural significance of a city, especially during HREs: as tourism destinations for activities (vi), as a traditional landmark at risk of losses (iii), as a collection of grandiose buildings (iv) and masterworks (i), as a representation of cultural influences (ii), and as a scenery spot (vii) in addition to being a cultural heritage.

Among the detected emotions and sentiments $\mathcal{Y}^{\mathrm{EMS}}$ possibly expressed in the tweets, the neutral ones are consistently dominant. In both cases, 'joy' and 'sadness' followed as equally sub-dominant explicit emotions, respectively pointing to the general sharing behaviour of people in an everyday context and the triggered sorrow after knowing the radical HREs. 'Anger' was consistently the 3rd most expressed emotion, although less significant in Notre-Dame (roughly 1/3 of sadness) than in Venice (more than 1/2 of sadness). 'Anger' as a main emotion being expressed is quite reasonable, as people could start looking for the actors to blame immediately (e.g., a group of people, some politicians, or a costly infrastructure). 'Fear' and 'surprise' were both detected in Notre-Dame and Venice, but not significant in either case and 'disgust' was never detected as the main emotion of a tweet.

The detected key topics of interest $\mathcal{Z}^s$ can be grouped hierarchically within six themes:

- **Emotions** that were composed of explicit emotion expressions or repeatedly used certain emojis. This is the most significant topic cluster in both cases.
- **Heritage** that explicitly or implicitly mentioned certain heritage values or attributes, such as "spire", "rose window", "architectural monument", and "artefact" in Notre-Dame, and "Venetians" and "holiness" in Venice.
- **Incidence** that reported the development and severeness of the event, such as "fire", "burn", "collapsed spire", and "destroyed ashes" in Notre-Dame, and "tide", "high water", "flood", "climate change" in Venice.
- **Actions** that either reflected on who and what to blame, such as "MOSE"[19] in Venice, or called for further actions as monetary and emotional supports, such as "help", "donation", "rebuilt", "reconstruct", "laser scanner", and "local management" in Notre-Dame, and "receive support" and "help Venice" in Venice.
- **Other Sites** that extensively mentioned another associated and/or unrelated place or person, such as "Louvre", "Victor Hugo", "Eiffel Tower", "Vatican", and "national Museum" in Notre-Dame, and "Biennale" and "Venice Beach" in Venice.
- **Politics** that referred to a politician, a political party, a movement, or a celebrity that can be relevant, such as "Emmanuel Macron", "elected officials", "yellow vest" and "Henri Pinault" in Notre-Dame, while no politics-related topics seemed related to Venice flood.

Moreover, even though the conventional practice of topic modelling using BERTopic would disregard the remaining documents that cannot be clustered into any existing topics, it is found in this study that the keywords generated from such a 'noise' topic had a clear connection to heritage management. For example, "heritage" and "San Marco" respectively appeared in the 'noise' topic of Notre-Dame and Venice. Therefore, it is kept and renamed as **Base**. The main themes detected here clearly resemble the repeated topic schemes in disaster management literature mentioned in Section 2.2. A full list of keywords for each sub-topic within the six themes can be found in Appendix B.

[19] MOdulo Sperimentale Elettromeccanico in Italian, literally translated as 'Experimental Electromechanical Module' https://en.wikipedia.org/wiki/MOSE, accessed 16 May 2023.

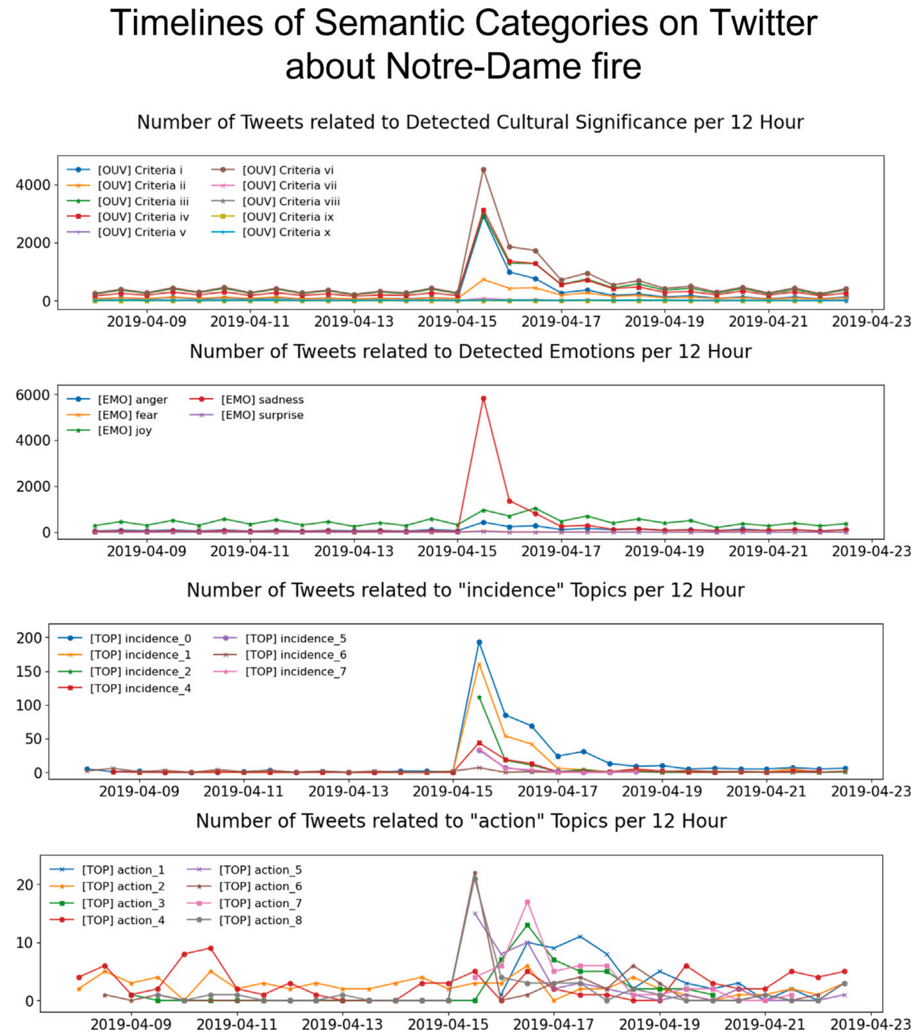

Fig. 8. A selection of timelines showing the temporal development of semantics along the HREs in Notre-Dame fire.

5.3. The spatiotemporal dynamics of semantics

The temporal development of the detected semantics, i.e., cultural significance $\mathcal{Y}^{\mathrm{OUV}}$, emotions $\mathcal{Y}^{\mathrm{EMS}}$, and key topics $\mathcal{Y}^{\mathrm{TOP}}$ along with the HREs can be inspected with timelines. A selection of highly relevant information is visualized in Figs. 8 and 9. The full collection of timelines with all detected semantic topics can be found in Appendix B. Different from Fig. 2 with hourly aggregation, the tweets are counted every 12 h in Figs. 8 and 9 for general temporal trends to emerge. Therefore, they can be treated as both the smoothed (with a longer time window) and factorized (as subsets of entire tweets) versions of the general timeline. The timelines demonstrate similar patterns: the tweets classified as related to certain semantics remained at a low level until the HREs happened, when the intensity rose to a very high level for a short period; afterwards the intensity drew back to normal. This pattern is more obvious in Notre-Dame fire as the contrast of intensity was higher.

For both cases, the cultural significance that rose the most during HREs were criterion (vi) about people's association, criterion (iii) about testimony, and criterion (iv) about architectural typology, as discussed in Section 5.2. Yet for Notre-Dame, discussions concerning criterion (i) about masterpieces also increased significantly, since people cared about the architectural monument (e.g., the spire and the rose window) and the important artefacts that could be destroyed by the fire, which were less worrisome in the Venice flood.

The emotions of sadness and anger were both triggered to rise. Sadness in Notre-Dame became 100-fold and got extremely dominant during the HREs, while in Venice the sadness became only 10-fold, both of which dropped to about 2- to 3-fold of baseline after HREs. Even though not as significant as sadness, the anger also remained higher since HREs happened compared to the calm baseline. Interestingly, the dominant emotion of joy before HREs also remained at a moderate level throughout the timelines and was back to dominance in the last days. This could be simply a result of a higher amount of posted tweets.

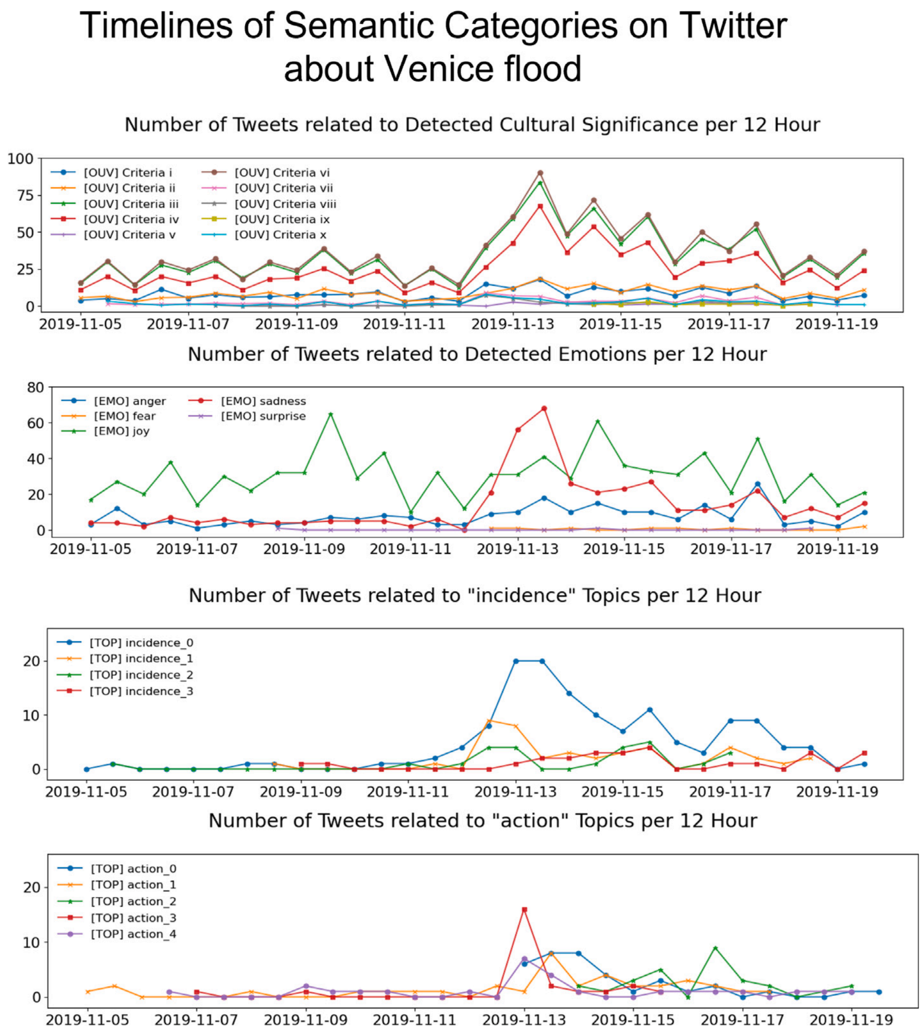

Fig. 9. A selection of timelines showing the temporal development of semantics along the HREs in Venice flood.

The majority of discussions describing the incidence and actions emerged on the same day HREs happened. In Notre-Dame, the descriptions of the fire in the cathedral (incidences 1, 2, 5) and the courage of firefighters (incidence 4) mainly appeared on 15 and 16 April UTC and diminished quickly afterwards, while incidence 0 mentioning the historic and symbolic meanings of Notre-Dame had another wave on 17 April, probably corresponding to the action 1 about the donations of French billionaires to rebuild the destroyed parts. Similarly, actions about rebuilding the Notre-Dame (action 5), the collapsed arrow in an identically modernized version (action 6) and other facilities with the help of local management (action 8) already existed immediately after the fire started, but were brought back on 16–17 April when donations were made (action 1, 3). Remarkably, on 17 April, another discussion went dominant (action 7) mentioning the late Belgian art historian Andrew Tallon and his work of using 3D laser scanning to build a digital model of Notre-Dame, as a prosperous source for restoration.[20] General actions such as thinking (action 4) and appraising (action 2) did not demonstrate a clear temporal pattern related to the fire.

In Venice, on the other hand, the most dominant description of the incidence as the worst flooding in 50 years (incidence 0) extended to a few days after the starting point on 13 November, probably because the topic also involved climate change and global warming as the hypothesized cause. In the later days of the flooding, a specific topic emerged reporting the damaged books by the flood in Bertoni bookshop located in San Marco (incidence 4). From the first days of the flooding, the MOSE project was mentioned (action 0, 3) and criticized as a failure costing billions of euros. Interestingly, starting on 14 November and reaching its climax on 16 November, an online campaign to save Venice by donating one euro for each selfie made was initiated by the Comune di Venezia (action 2), following the discussion of support made by companies (action 1).

[20] https://www.vassar.edu/stories/2019/190417-notre-dame-andrew-tallon.html, accessed 22 May 2023.

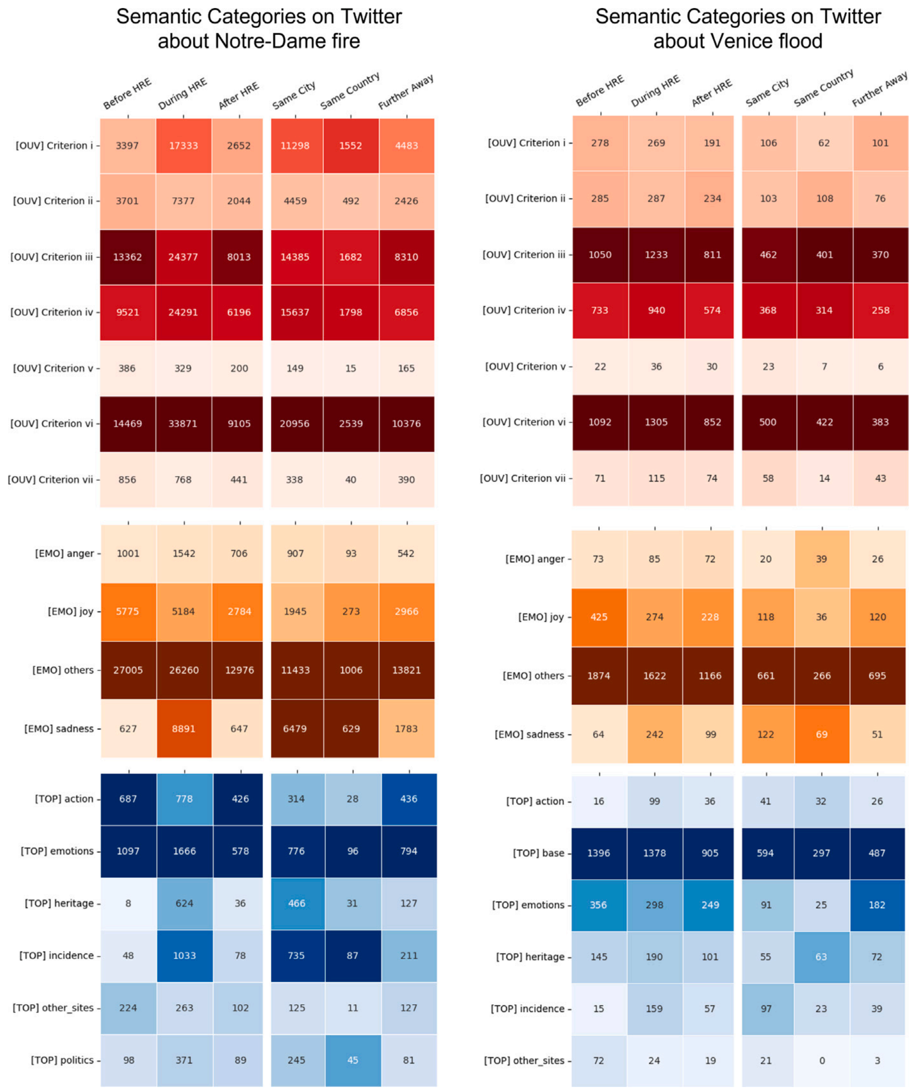

Fig. 10. The distribution of Tweets categorized with pseudo-labels of top-3 OUV selection criteria, detected emotions, and key topics under each theme for different periods and localities during HREs in Notre-Dame and Venice. The tweets belonging to the semantic type (row) and the period/locality type (column) are counted as a value matrix in the cell, while the heatmap is coloured using the column-normalized matrix denoting the proportions. The row-sum of cells with different localities equals the cell 'during HREs' of the same row. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Aggregating the tweets under each semantic type for different periods (before, during, and after HREs) and localities (same city, same country, and further away), the distributions are visualized in Fig. 10. The categories that are too over-representative (the 'base' topic) or too scarce (the OUV selection criteria viii-x, and the emotions fear and surprise) are omitted. The distributions varied with different periods and localities. For example, sadness emotion and incidence topic were consistently higher during HREs than before and after, and more posted (proportionally, not necessarily numerically) in the same city than further away. However, the two case studies also demonstrated different spatiotemporal patterns. In Notre-Dame fire, tweets concerning OUV criterion (i) about masterpieces were detected during HREs posted by people from France; people from Paris and France expressed extensively their sorrow while reporting the fire as an incidence and possible damage to it, yet people from further away tried to suggest and/or take various actions to help. In Venice flood, on the contrary, anger (probably associated with MOSE), action-related, and heritage-related discussions were detected in Italy, while emotions- and/or emoji-related tweets were posted more further away. Such observations are further justified with Chi-square Contingency tests,[21] a non-parametric version of two-way ANOVA for categorical variables, as reported in Table 2. Even though all the independent Chi-square tests showed a significant difference in distributions across different

[21] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.chi2_contingency.html, accessed 20 May 2022.

Table 2
Independent Chi-square tests on the distributions of semantic labels across different periods and localities. The effect size Cramer's V is calculated as $V = \sqrt{\chi^2/(n \times df*)}$ following Gravetter et al. [106], where $df*$ is the minimum of the number of rows or columns minus 1 (consistently $df* = 2$ in this case).

Statistics		Notre-Dame fire				Venice flood			
		χ^2	n	df	V	χ^2	n	df	V
OUV	Periods	3639.9***	182,689	12	.100[b]	24.7*	10,482	12	.034[a]
	Localities	646.2***	108,346	12	.055[a]	50.2***	4 185	12	.077[b]
Emotions	Periods	8584.0***	93,398	6	.214[c]	151.8***	6 224	6	.110[b]
	Localities	3245.8***	41,877	6	.197[b]	95.2***	2 223	6	.146[b]
Topics	Periods	1209.6***	8 206	10	.271[c]	221.7***	5 515	10	.142[b]
	Localities	470.7***	4 735	10	.223[c]	154.6***	2 148	10	.190[b]

* $p < .05$.
*** $p < .001$.
[a] Very small effect size.
[b] Small effect size.
[c] Medium effect size.

periods and localities with $p < .01$, the effect size — Cramer's V in OUV tests was generally small. In Notre-Dame fire, the differences in the distributions of emotions and topics mostly had a medium effect size, while for the Venice flood, the effect sizes were always small. This complex spatiotemporal dynamic of distributions for each semantic type invites further investigations.

6. Discussion

6.1. Indications for heritage management under disaster risks

Even though obtained empirically with unsupervised methods, the six detected hierarchical themes of key topics in Section 5.2 also conceptually overlapped with the other two semantics, i.e., emotions and cultural significance, at a different level of details. This also justifies the choice in the beginning of analysing tweets during HREs from these perspectives. As already mentioned in Section 4.3.3, the themes detected largely resonate the existing schemes [61,65]. This suggest that cultural heritage does integrate well with the general approaches of disaster risk management albeit not conventionally "taken into account in global statistics" [9]. The main difference comes with a specific focus on cultural significance with the framework of OUV selection criteria involving heritage values and attributes, which are the novel additions offered by this study.

The outcomes of this paper yield two meaningful observations for heritage management under disaster risks:

- People will extensively express sadness and relentlessly share information about the damage during HREs. HREs will trigger discussions online and involve concerned people from far beyond, transcending geographical boundaries. Such well-known knowledge and "common sense" are confirmed empirically.
- Previously less-known and/or surprising information also emerged. This includes the criticism of the MOSE project in Venice, the campaign of "#oneeuroforoneselfie" by the municipality, the rediscovery of the work by Professor Andrew Tallon for Notre-Dame Cathedral, and the volume of tweets expressing joy and anger during HREs in both case studies.

The confirmation of well-known knowledge is meaningful as it can support the planning decisions made based on experience and heuristics on what to prioritize [38,65]. It also shows the necessity for expanding the definition of heritage community suggested in Faro Convention [107] in the time of radical events, by including the temporally-founded communities bonded by the HREs into the scope, as determining who should contribute to the discourse is a crucial step in disaster risk management [4,9,37,62]. Moreover, the discovery of less-known information is even more valuable. On the one hand, not all scholars and practitioners studying an HRE are acquainted with the heritage property in depth, especially in global collaborations (with geospatial and sociopolitical distance) and retrospective historic investigations (with temporal and cultural distance). On the other hand, even for local administrators and sophisticated scholars studying a heritage property for years, specific aspects of knowledge can still be overlooked with information bubbles and confirmation bias [108,109].

Even though the international and national doctrines on cultural heritage management under disaster risks still lack the mentioning of social media and artificial intelligence [8,9,11,12], they offer substantial potentials and opportunities [38,65]. Requirements and strategies brought up far before the digital age by Stovel [6] can be realized more easily, such as enhancing appropriate documentation systems, mobilizing media as positive force in shaping public perceptions, increasing the awareness and sensitivity of stakeholders, providing generalizable recommendations relevant for all sites of cultural heritage values, reviewing large range of alternatives, etc. These new roles social media could play are also gradually getting noticed by practitioners together with the new visions towards the List of World Heritage in Danger [22]. The workflow demonstrated in this paper provides the possibility for users to acquire information at a large scale, being an effective and transferable knowledge documentation tool useful for inclusive heritage management and planning under constant climatic and environmental changes including the impacts from disasters, as suggested by the Recommendation on the Historic Urban Landscape [10,23]. The data collection periods of before, during, and after the HREs correspond to the phases of preparedness, response, and recovery for disaster management, when potential activities such as funding, emergency response, and awareness can be taken [6,61,62]. Interestingly, from a retrospective view, as a case study with “lengthy process” promoted already in [9] to mitigate disaster risk, MOSE was not yet fully completed and put in use in 2019, and it managed to prevent Venice from an even larger flooding in 2022.[22] In this case, the proposed framework can also help engineers and technology historians reveal the dynamics of public reactions and risk perceptions towards a major project in the domain of environmental protection, recovery, and management.

The tools proposed in this paper can be understood as an “observatory” of specific heritage properties, which can be eventually turned into a dashboard or “thermometer” with “fully automated and near real-time event detection and tracking system” [27] to monitor the reactions and social sentiments of the public, in pursuit of social inclusion and consensus building [4,10,21]. This is especially useful when there lacks sufficient reporting of disaster risks from the administrative level. Social media gives chances to reveal more knowledge than what is told in the periodic reporting exercises, i.e., the State of Conservation reports, and the Reactive Monitoring missions [110,111]. As such, experience learnt from best practices and even failures can be disseminated to similar cultural heritage properties world wide, leading to more effective disaster risk reduction strategies, plans, crisis communication, and operations [9,10,16,22,61,63].

When commenting on the usage of online media during a radical event in the digital era, Garduño Freeman and Gonzalez Zarandona [17] brought up the examples of the Notre Dame fire and Palmyra destruction. Whereas the search volume on Google Ads increased 60-fold in response to the former event, it only increased seven and a half times for the latter. Garduño Freeman and Gonzalez Zarandona [17] criticized that this seemed to suggest that “one site was mourned by more people than the other, so much so that it has created the impression of a **Notre-Dame effect**”. They further argued that the so-called Notre-Dame effect and the broader concept of “mediatisation of heritage” entailed digital colonialism, challenging the equality and equity of WH co-existing in the same list. Even though the case studies in this paper, Notre-Dame and Venice, both come from the “Canon” of European history, a similar effect of one type of HREs raising more attention than the other emerged. The consistencies, nuances, and differences between the two cases suggest that there might be some general rules behind people’s online reactions in HREs, which are applied at a different level and adapted to the specificity of a disastrous radical event. Therefore, extra consciousness is needed while interpreting the results with future applications in other case studies of HREs globally. Especially, the cultural significance of a WH property should not be over-simplified as a set of “valuable linguistic metonyms” (keywords) targeted at only specific groups of audiences [17].

6.2. Limitations and future studies

The data collected in the study was restricted to the ones that were either initially given a geo-location or directly connected to the tweets with geo-locations, which automatically limited the scope of the study since only a small proportion of tweets are geo-tagged [24]. Future studies could lift this restriction and collect a larger initial dataset and query for tweets both directly and/or indirectly (with a distance of 2–3 network steps) responding to and being responded to by the seed tweets. Afterwards, Named-Entity Recognition [112] could be used to infer the geo-locations if not explicitly given [113]. As such, the modules of constructing graphs and generating semantic labels are still valid, whereas a more comprehensive view of the conversation dynamics could be obtained, albeit less focused on the spatial aspects. Only collecting data from Twitter users with a few languages also limit the pool of information, which may lead to biases and exclusiveness. Voices from people not using social media or reacting in other languages on other platforms (such as Instagram, Facebook, Reddit, Tiktok or the more localized WeChat and RED) are *de facto* over-looked. Since the proposed methodological workflow is modular, analyses could be easily repeated in future studies after including diverse data sources of specific interests. Furthermore, results from social media analyses can be also paired with qualitative studies involving policy analysis, interviews, and/or focus groups, in order to gain more comprehensive insights from broader stakeholders, amplifying the validity of similar studies through “triangulation” of multiple sources [9]. Under the same argument of modularity, classification frameworks and models of semantic information can also be replaced and/or enhanced with other alternative categories, e.g., the eight heritage values (historic, aesthetic, social, political, etc.) proposed and continuously applied by Pereira Roders [82] and the schemes summarized in [61].

Traditional topic modelling algorithms such as LDA are known as unstable against different configurations, hard to reproduce, and work badly with short texts such as tweets [114]. The use of BERTopic resolves the issue and makes it possible to obtain clear topics and fine-grained tweet-level predictions. Depending on potential questions of interest, the tweets could also be merged at the level of users, communities, interest groups, cities, countries, and/or spatiotemporal clusters. Moreover, human experts are always needed to control the quality of topics and select the relevant ones for further interpretation. Several decisions were arbitrarily made while demonstrating the methodology, such as the filtering thresholds for cultural significance in Section 4.3.1 and topics in Section 4.3.3, the model selection in Section 4.3.2, and the hyperparameters for BERTopic in Appendix A. Further experimentation with these decisions using either qualitative inspections or quantitative metrics could potentially test the robustness, validity, and reliability of the results. Among others, the clustering quality measures could include the Davies–Bouldin Index, the Dunn Index, the Silhouette Efficient [42], clustering consolidation and discrimination [45], and Strength Index [28]. Different from the conventional event detection studies [24,27,28], where spatiotemporal clustering algorithms are first used to find significant clusters before semantically describing them, this paper skipped the spatiotemporal clustering. Instead, tweets were clustered implicitly with their semantics with topic modelling [104], as a pragmatic choice. However, including an additional step of spatiotemporal clustering

[22] News article Marea a Venezia. Il Mose salva la città, l’acqua tocca 204 centimetri, accessed 26 May 2023.

either before or after the topic modelling could give another layer of interpretation. The answered questions would therefore become "What are the expressed emotions and main semantic topics being discussed within each cluster that is significantly distinguishable by its spatiotemporal density", which could also be an interesting topic for future studies.

Furthermore, starting from the collected dataset and conducted exploratory analyses, many more interesting questions in the fields of heritage studies, urban studies, social sciences, disaster risk management, computer science, and Geo-AI research could be answered. By repeating the procedure in other HREs and/or other types of disasters happening to heritage properties with different geopolitical and cultural backgrounds in different years [27,37,39], general rules of online interaction discussed in 6.1 could be verified, resulting in a handbook for heritage managers on how to act and react on social media during events to reduce further secondary risks [15]. By digging into the semantic development embedded in the conversational graph structure, the mechanisms of information spreading, stance changing, and interactional framing could be further revealed [18,60,115], closing the loop of combining all four dimensions of social media data [61] as mentioned in Section 2.2. The time zones, language, and social interests of users can all be used as grouping variables to describe and explain the spatiotemporal patterns of posting behaviour, the development of semantics, and the interaction mechanisms behind them. Such mechanisms should be generalizable across fields beyond the scope of heritage, being able to explain debates on other societal issues triggering public interactions, such as sustainability actions, climate change campaigns, and global pandemics [14,59,116]. Only textual information was analysed for the semantics, yet a multi-modal representation including images, memes, audio, and videos can add other contradictory or complementary information [20,117,118]. The emojis used in tweets were not thoroughly investigated in this paper, similar operations as in [35] could also be conducted to compare the change of emoji usage before, during, and after the HREs. The results of this study could be combined with other relevant works collecting information using social media in Notre-Dame fire [53,55] and Venice flood [69,119], to construct a multi-layer digital archive concerning the event. This, in turn, can eventually lead to a better, more socially inclusive, and informative risk management plan for cultural heritage [5,9,30,120–122].

7. Conclusions

This paper presents a methodological framework to investigate the collective behaviour on social media when radical Heritage-related events (HREs) happen. It applies a few pre-trained natural language processing models to obtain pseudo-labels of tweets in the time of HREs for their semantic meanings including conveyed cultural significance, expressed emotions, and discussed topics. The conversational sequences and the spatiotemporal contexts are modelled in graph structures. Two case studies that both happened in 2019, the fire in Notre-Dame and the flood in Venice, are used as demonstrative examples to showcase the framework. Exploratory data analysis and statistical tests are conducted to describe the spatiotemporal dynamics of the reactions of the online public from the same city, the same country, and far beyond within the periods before, during, and after HREs. Results show that the online discussions went far beyond the local heritage community and triggered vivid expressions of emotions and action proposals globally, even though the discussion waves drew back quickly afterwards. The methodological framework can be also applied to other similar disastrous events happening to heritage globally, facilitating inclusive heritage management processes and reducing risks against cultural significance. It functions as an information gathering and eventually, a knowledge documentation tool to confirm known facts and discover new ones.

CRediT authorship contribution statement

Nan Bai: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Pirouz Nourian:** Writing – review & editing, Validation, Supervision, Methodology, Formal analysis, Conceptualization. **Tao Cheng:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Conceptualization. **Ana Pereira Roders:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The code of this work is available on 4TUResearchDataRepository.

Acknowledgement

The presented study is within the framework of the Heriland-Consortium. HERILAND is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813883.

Appendix A. Model implementation details

To obtain the key topics as semantic information of tweets, the BERTopic python library was used [104]. The inputs of the topic models were the translated and normalized tweets, as mentioned in Section 3.2, where each tweet was regarded as an individual document. Within the six main modules of the BERTopic library, the configurations were as follows:

1. The Sentence-Transformer model "all-MiniLM-L6-v2" [123] trained in English was used as the embedding model for the input texts.
2. The UMAP (Uniform Manifold Approximation and Projection) model with default parameter configurations was used as the dimensionality reduction algorithm, where the option of low memory was selected to prevent large datasets such as Notre-Dame fire from running out of memory.
3. The default HDBSCAN model was used to cluster the vector representations with reduced dimensions with a minimum topic size of 45 in the case of the Notre-Dame fire and 25 in the Venice flood.
4. For each cluster, the CountVectorizer tool from the Scikit-Learn Python library was used as the vectorizer model to obtain a bag-of-words matrix, where both single words and 2-grams (two consecutive words) were counted and the stop words lists provided by NLTK Python library for English and the local language (French or Italian) were excluded. Note that the stop words were only excluded here at a later stage for generating the verbal description (representation) of the cluster, but not before the sentence embedding step, since Transformer-based BERT models prefer to view words in their semantic contexts.
5. The adjusted version of TF–IDF, the c-TF–IDF (class-Term Frequency–Inverse Document Frequency) was used on the level of clusters by combining the bad-of-words matrices of all tweets belonging to each cluster. Specifically, the importance of very frequent words after removing the stop words was further reduced by taking the square root of all term frequencies. The initially obtained clusters were then automatically reduced by another round of HDBSCAN clustering on the c-TF–IDF cluster representations, resulting in the final detected topics.
6. To improve the quality of the obtained topic representations, the algorithm of Maximal Marginal Relevance was used to decrease the redundancy of keywords and increase the diversity of keywords for each topic.

The matrix showing the probabilities of each tweet belonging to each obtained topic was calculated and saved. All the topics $\mathcal{Z}$ together with their keywords representation were checked manually to select the ones that might be relevant and interesting for this research $\mathcal{Z}^s$, and categorized into six themes: emotions (emoji), heritage, incidence, actions, other sites, and politics, as already described in Section 5.2. All the other topics that were not selected were ignored for further analysis in this research.

Appendix B. Extended results

B.1. List of interesting topics for the Notre-dame fire

Fig. B.1 shows the complete timelines of all semantic categories of cultural significance, emotions, and key topics detected in the Notre-Dame fire dataset. A selection has been previously illustrated in Fig. 8. The keywords associated with each detected interesting topic under each theme are listed below, note the emojis are transformed into verbal descriptions:

- **Base**
  - [TOP] base 0: church, heritage, dame paris, dame cathedral, via, burning, dame fire, rebuild, notredamedeparis, may
- **Emotions**
  - [TOP] emotions 0: face_with_tears_of_joy face_with_tears_of_joy, user face_with_tears_of_joy, face_with_tears_of_joy httpurl, ça face_with_tears_of_joy, plus face_with_tears_of_joy, aussi face_with_tears_of_joy, oui face_with_tears_of_joy, face_with_tears_of_joy vtep, grave face_with_tears_of_joy, know face_with_tears_of_joy
  - [TOP] emotions 1: notredame notredame, httpurl notredame, notredame sad, crying notredame, httpurl sad, believe notredame, cry notredame, notredamedeparis notredame, attack notredame, awful notredame
  - [TOP] emotions 2: loudly_crying_face loudly_crying_face, loudly_crying_face httpurl, loudly_crying_face red_heart, loudly_crying_face dame, loudly_crying_face face_with_tears_of_joy, loudly_crying_face crying_face, crying_face loudly_crying_face, loudly_crying_face broken_heart, loudly_crying_face paris, baby loudly_crying_face
  - [TOP] emotions 3: eyes smiling_face_with_heart, smiling_face_with_3_hearts, smiling_face_with_3_hearts smiling_face_with_3_hearts, smiling_face_with_3_hearts user, beaming_face_with_smiling_eyes beaming_face_with_smiling_eyes, smiling_face_with_3_hearts httpurl, user beaming_face_with_smiling_eyes, smiling_face_with_sunglasses smiling_face_with_sunglasses, smiling_face_with_smiling_eyes smiling_face_with_smiling_eyes, user smiling_face_with_smiling_eyes
  - [TOP] emotions 4: red_heart red_heart, yellow_heart, love red_heart, user blue_heart, green_heart, purple_heart purple_heart, red_heart notredame, blue_heart httpurl, red_heart thank, merci red_heart
  - [TOP] emotions 5: broken_heart notredame, loudly_crying_face notredame, crying_face notredame, notredame crying_face, notredame loudly_crying_face, face_screaming_in_fear loudly_crying_face, sad_but_relieved_face notredame, face_screaming_in_fear notredame, crying_face notredamedeparis, broken_heart loudly_crying_face

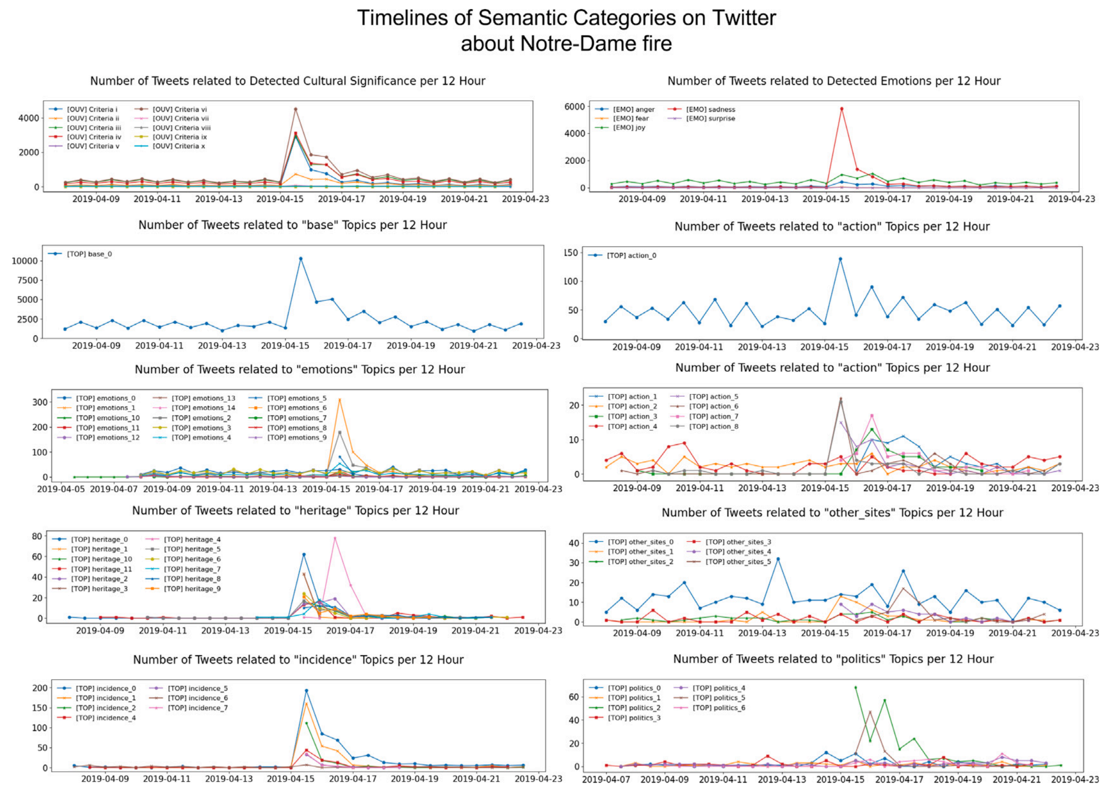

Fig. B.1. The complete timelines showing the temporal development of semantics along the HREs in Notre-Dame fire.

- [TOP] emotions 6: grinning_face_with_sweat grinning_face_with_sweat, thinking_face grinning_face_with_sweat, merci grinning_face_with_sweat, grinning_face_with_sweat ouf, jew optimistic, way grinning_face_with_sweat, grimacing_face grinning_face_with_sweat, go grinning_face_with_sweat, grinning_face_with_sweat red_heart, grinning_face_with_sweat virgintonic
- [TOP] emotions 7: face_screaming_in_fear face_screaming_in_fear, tired_face face_screaming_in_fear, face_screaming_in_fear face_with_monocle, face_screaming_in_fear juvaja, understand face_screaming_in_fear, face_screaming_in_fear loudly_crying_face, flushed_face face_screaming_in_fear, face_screaming_in_fear cold_face, face_screaming_in_fear heritage, speechless face_screaming_in_fear
- [TOP] emotions 8: anxious_face_with_sweat, anxious_face_with_sweat anxious_face_with_sweat, anxious_face_with_sweat user, user anxious_face_with_sweat, httpurl anxious_face_with_sweat, anxious_face_with_sweat loudly_crying_face, anxious_face_with_sweat pensive_face, non anxious_face_with_sweat, hot_face anxious_face_with_sweat, 징글징글 anxious_face_with_sweat
- [TOP] emotions 9: face_vomiting face_vomiting, nauseated_face face_vomiting, nauseated_face nauseated_face, face_vomiting angry_face, angry_face face_vomiting, user face_vomiting, face_vomiting nauseated_face, reading face_vomiting, islamophobia like, innocuous survey
- [TOP] emotions 10: face_screaming_in_fear, httpurl face_screaming_in_fear, face_screaming_in_fear face_screaming_in_fear, face_screaming_in_fear crying_face, user face_screaming_in_fear, grinning_face face_screaming_in_fear, omg face_screaming_in_fear, words face_screaming_in_fear, quality diversity, provider face_screaming_in_fear
- [TOP] emotions 11: broken_heart user, broken_heart, broken_heart broken_heart, heartbroken broken_heart, confounded_face broken_heart, sad broken_heart, miskina broken_heart, pain meditation, misha tweet, mothers suicides
- [TOP] emotions 12: hug ganchita, hugs user, giant hug, ganchita thank, need hug, hug great, hug tds, hug tilda, hug viet, hug jesus
- [TOP] emotions 13: pleading_face, pleading_face pleading_face, wsh pleading_face, pleading_face damn, like pleading_face, pleading_face pensive_face, st pleading_face, thank pleading_face, time pleading_face, jsuis pleading_face
- [TOP] emotions 14: shocked user, surprise user, user shock, shock user, know shocked, shock, recal box, jui shocked, part surprise, policeman charge

- **Heritage**
  - [TOP] heritage 0: notredame paris, paris notredame, httpurl notredame, paris httpurl, notredame symbol, france notredame, day france, symbol france, httpurl sad, today paris
  - [TOP] heritage 1: spire collapsed, spire collapses, spire collapse, fire spire, cathedral collapses, paris collapsed, collapsed fire, collapse dame, spire cathedral, spire roof
  - [TOP] heritage 2: rose window, rose windows, stained glass, glass windows, windows survived, window dame, window spared, rosettes, survived fire, three rose
  - [TOP] heritage 3: notredame spire, collapsed notredame, spire fell, two towers, moment notredame, roof notredame, towers fire, collapse spire, roof collapsed, towers notredame
  - [TOP] heritage 4: important artefacts, saved brave, fire positive, full important, braveheroes, positive signs, firefighters braveheroes, rebuilt restored, iconic building, notredame rebuilt
  - [TOP] heritage 5: monuments, historical monuments, heritage professionals, monument like, monument user, people built, magnificent monument, conservators archaeologists, avoid facelift, historians kills
  - [TOP] heritage 6: cathédrale dame, cathédrale, user cathedral, paris cathedral, cathedral httpurl, broken_heart dame, broken_heart cathédrale, cathedral dame, dame red_heart, black_heart church
  - [TOP] heritage 7: cross, palm sunday, holyweek, notredame cross, cross stands, cross christ, arms cross, latin_cross, httpurl jesus, joseph
  - [TOP] heritage 8: discussing historic, significance dame, homage dame, stone paper, place saint, pays homage, witness httpurl, paris work, dame paris, marseille leans
  - [TOP] heritage 9: church notredame, notredame catholic, catholics, church building, catholic church, notredame owned, notredame much, place worship, catholic religion, notredame church
  - [TOP] heritage 10: fire notre_dame, fdny, fire symbol, dame fire, historic houses, invaluable places, legacy fire, library mention, life dozens, history cry
  - [TOP] heritage 11: paris cathedral, user dame, cathedral httpurl, garde photo, user cathédrale, dame paris, cathédrale dame, millefeuille dame, chapelle onze, kapellekerk
- **Incidence**
  - [TOP] incidence 0: dame user, dame dame, dame fire, dame burning, dame burns, like dame, gothic, dame symbol, history dame, dame cathedral
  - [TOP] incidence 1: paris cathedral, cathedral paris, cathedral dame, cathedral fire, user cathedral, dame paris, fire paris, fire breaks, cathedral notredame, httpurl cathedral
  - [TOP] incidence 2: notredame fire, fire notredame, notredame burning, notredame notredamecathedralfire, fire notredamedeparis, httpurl notredamecathedralfire, fire paris, notredame paris, flames notredame, notredamecathedralfire notredame
  - [TOP] incidence 4: courage firefighters, firefighters notredame, notredame firefighters, congratulations firefighters, firefighters mobilized, hope firefighters, heroes, fire firefighters, yubari, dear firefighters
  - [TOP] incidence 5: paris fire, fire paris, dame paris, depths laments, laments stéphane, dame fire, ee, france affected, paris homework, video fire
  - [TOP] incidence 6: fire user, fire fire, user fire, sub rogue, gros fire, ignites user, spontaneously ignites, smart plug, anything fire, firecatchesfire user
  - [TOP] incidence 7: reduced ashes, sad fire, ashes fire, cathedral burnt, cathedral burning, cathedral burned, cathedral fire, france stfu, flames survided, fire mum
- **Actions**
  - [TOP] action 0: helped turn, accounts helped, verified accounts, user trndnl, topic user, accounts, love user, user merci, know user, awful user
  - [TOP] action 1: donations, donate, rebuild dame, donations dame, donated, donating, french billionaires, money rebuild, millionaires, pledges
  - [TOP] action 2: thumbs_up thumbs_up, user thumbs_up, thumbs_down, thumbs_down thumbs_down, pretty scallop, ideas rainbow, httpurl thumbs_up, kiss_mark like, thumbs_up collection, thumbs_down httpurl
  - [TOP] action 3: donations, donations notredame, taxes taxes, reconstruction notredame, notredame donations, million euros, notredamedesriches, donate notredame, millionaires, notredame billion
  - [TOP] action 4: thinking_face thinking_face, rapport thinking_face, turn thinking_face, investigation thinking_face, talking thinking_face, answer thinking_face, paris thinking_face, eyes thinking_face, coincidence thinking_face, something thinking_face
  - [TOP] action 5: rebuild notredame, notredame years, notredame rebuilt, deadline rebuild, accomplished five, rebuilding notredame, reconstruction notredame, saying rebuild, notredame rebuild, rebuilt years
  - [TOP] action 6: arrow notredame, arrow collapsed, arrow fall, rebuild arrow, new arrow, identically modernize, arrow magnificent, eyes arrow, httpurl arrows, arrow adapted

- [TOP] action 7: 3d, andrew tallon, historian laser, helping rebuild, scans dame, architectural historian, laser scanners, used lasers, worked laser, historian andrew
- [TOP] action 8: user rebuilt, less complexity, local management, square construction, include gdf, lift petticoats, indeed facilities, irreplaceable ok, irl workflow, maintained cleaning

• **Other Sites**

- [TOP] other sites 0: louvre, httpurl paris, île france, paris île, picasso, musée, streetphotography, paris paris, seine river, gallery
- [TOP] other sites 1: victor hugo, hugo dame, 1831, hugo hunchback, hugo novel, hugo wrote, dame victor, empty skeleton, miserables, novel dame
- [TOP] other sites 2: 875 875, priests, xvi, pedophilia, churches france, pope benedict, churches attacked, pedophilia church, france vandalized, 875 churches
- [TOP] other sites 3: eiffel, eiffel tower, tower paris, floor eiffel, tower every, see eiffel, francissantamaria eiffel, restaurant eiffel, towereiffel, top eiffel
- [TOP] other sites 4: vatican, catholic church, say vatican, church donors, richest institutions, vatican sitting, much vatican, vatican give, church afford, user vatican
- [TOP] other sites 5: national museum, brazilian billionaire, 88 million, billionaire donated, donated 10, brazilians, brazilians donate, dame give, brazilian woman, find brazilian

• **Politics**

- [TOP] politics 0: emmanuel macron, macron20h, user macron, macron elected, debate emmanuel, macron want, speech, macron speak, macron20h httpurl, president
- [TOP] politics 1: vote, politicians, senate, elected, electoral, president republic, elected officials, republic user, voters, prime minister
- [TOP] politics 2: emmanuel macron, president macron, french president, five, macron dame, macron promises, years httpurl, macron notredame, macron rebuild, rebuild cathedral
- [TOP] politics 3: algeria, sudan, rwanda, genocide, algerians, ottoman, egypt, tunisia, african hemicycles, arab world
- [TOP] politics 4: yellowvests, yellow vests, yellow vest, yellowvests paris, yellowvests acte23, 20 yellowvests, ultimatum2, protest, paris protest, yellowvests actexxii
- [TOP] politics 5: pinault family, henri pinault, million euros, arnault family, bernard arnault, donation, francois, renounces tax, paris pinault, billionaire
- [TOP] politics 6: yellow vests, vests dame, vest movement, paris yellow, vest protesters, rebel yellow, levavasseur calls, funds march, paris protests, vests protesting

B.2. List of interesting topics for the Venice flood

Fig. B.2 shows the complete timelines of all semantic categories of cultural significance, emotions, and key topics detected in the Venice flood dataset. A selection has been previously illustrated in Fig. 9. The keywords associated with each detected interesting topic under each theme are listed below, note the emojis are transformed into verbal descriptions:

• **Base**

- [TOP] base 0: water, italy, venice, see, marco, venezia, day, city, san marco, user venice

• **Emotions**

- [TOP] emotions 0: httpurl httpurl, user httpurl, httpurl user, httpurl fuck, httpurl love, httpurl oh, httpurl understand, httpurl new, httpurl god, httpurl excuse
- [TOP] emotions 1: face_with_tears_of_joy, face_with_tears_of_joy face_with_tears_of_joy, face_with_tears_of_joy user, user face_with_tears_of_joy, face_with_tears_of_joy httpurl, loudly_crying_face loudly_crying_face, loudly_crying_face user, loudly_crying_face face_with_tears_of_joy, face_with_rolling_eyes face_with_rolling_eyes, fearful_face fearful_face
- [TOP] emotions 2: red_heart red_heart, africa gem_stone, aristoflownetwork copyright, rè gem_stone, registered ã^{n}d, level zero, beating_heart rè, gem_stone beating_heart, hätè gem_stone, prohibited gem_stone
- [TOP] emotions 3: beaming_face_with_smiling_eyes, eyes smiling_face_with_heart, smiling_face_with_smiling_eyes smiling_face_with_smiling_eyes, smiling_face_with_3_hearts, beaming_face_with_smiling_eyes beaming_face_with_smiling_eyes, beaming_face_with_smiling_eyes user, beaming_face_with_smiling_eyes httpurl, smiling_face_with_3_hearts httpurl, grinning_face_with_smiling_eyes grinning_face_with_smiling_eyes, beaming_face_with_smiling_eyes red_heart
- [TOP] emotions 4: pleading_face, pleading_face httpurl, pleading_face pleading_face, pleading_face growing_heart, expressionless_face, face_with_monocle face_with_monocle, cazzie video, growing_heart httpurl, exploding_head exploding_head, blue_heart blue_heart

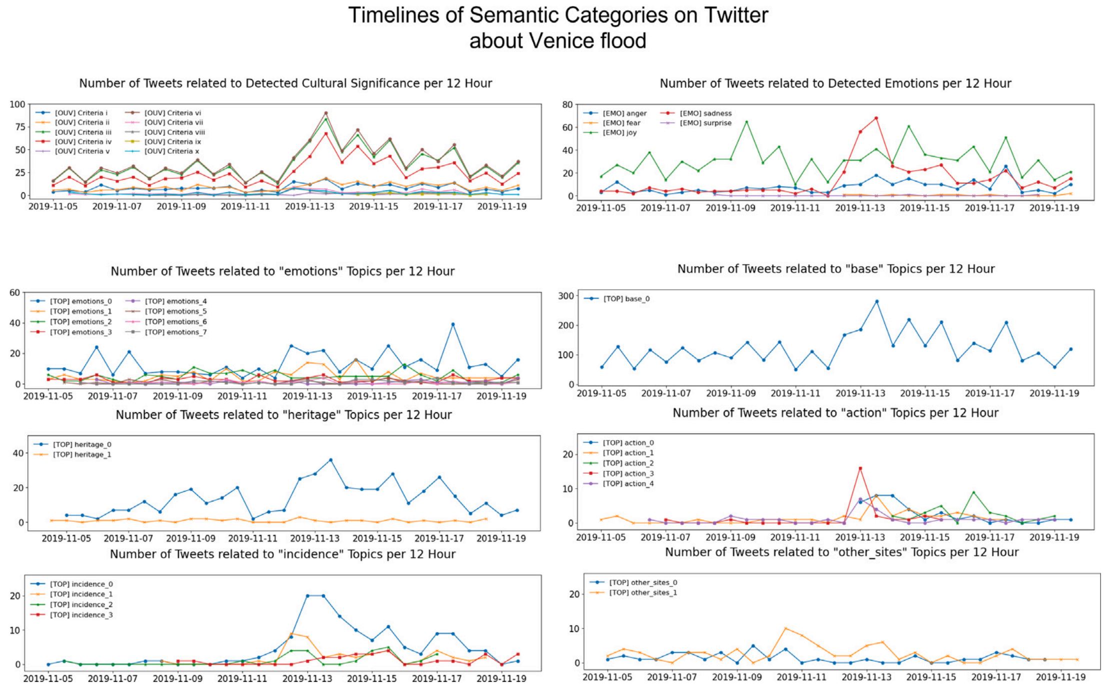

Fig. B.2. The complete timelines showing the temporal development of semantics along the HREs in Venice flood.

- [TOP] emotions 6: rolling_on_the_floor_laughing, rolling_on_the_floor_laughing rolling_on_the_floor_laughing, rolling_on_the_floor_laughing face_with_tears_of_joy, rolling_on_the_floor_laughing httpurl, face_with_tears_of_joy rolling_on_the_floor_laughing, venice rolling_on_the_floor_laughing, rolling_on_the_floor_laughing grinning_squinting_face, face_screaming_in_fear rolling_on_the_floor_laughing, rolling_on_the_floor_laughing tagadala7, face_savoring_food face_savoring_food
- [TOP] emotions 7: clapping_hands, clapping_hands clapping_hands, clapping_hands user, user clapping_hands, thumbs_up clapping_hands, clapping_hands flexed_biceps, clapping_hands ok_hand, clapping_hands top_arrow, gesture, clapping_hands httpurl

- **Heritage**
  - [TOP] heritage 0: salvini, venice italy, italia, league, venezia italia, venice, venetians, venice matera, venice acquaalta, venezia venice
  - [TOP] heritage 1: holiness, shamrock cherry_blossom, sins, shamrock, bright_button shamrock, cherry_blossom bright_button, allah, graduation, deceased, prayer
- **Incidence**
  - [TOP] incidence 0: climate, climate change, flooding, floods, flood, worst flooding, flooding venice, flooding 50, venice flooding, global warming
  - [TOP] incidence 1: high tide, highest tide, tide 50, centimeters, httpurl small_orange_diamond, flooded highest, city hit, exceptional tide, tide hits, hit highest
  - [TOP] incidence 2: sirens, siren, siren sounded, sirens sounded, four tones, sirens venice, alarm siren, httpurl sirens, water_wave water_wave, hear sirens
  - [TOP] incidence 3: bookstore, alta bookshop, library venice, bertoni bookshop, acquaalta bookstore, books destroyed, many books, disappointed_face disappointed_face, person_raising_hand person_raising_hand, calle
- **Actions**
  - [TOP] action 0: bribes, mose work, operation, project, veneto region, billion mose, venice mose, commissioners, euros spent, galan zaia
  - [TOP] action 1: venezia httpurl, httpurl venise, acquaaltaavenezia backhand_index_pointing_down, receives support, blame political, numbers people, read history, companies involved, venessiamia httpurl, unloading
  - [TOP] action 2: donate, magna, savevenice oneeuroforoneselfie, shareit helpvenice, helpvenice veneziaacquaalta, salviamovenezia comunedivenezia, euro could, oneeuroforoneselfie salviamovenezia, help city, million

- [TOP] action 3: sanservolo say, mose sanservolo, venezia mose, warning warning, acquaaltaavenezia httpurl, mose acquaalta, senator morra, impeachmenthearings fight_fight_against_cyber_violence, shit pile_of_poo, stikstofcrisis togetherforwonho
- [TOP] action 4: folded_hands, folded_hands folded_hands, venice folded_hands, user prayers, user folded_hands, writing_hand frasinliberta, person_raising_hand person_raising_hand, backhand_index_pointing_down backhand_index_pointing_down, speechless, palms_up_together

• **Other Sites**

- [TOP] other sites 0: biennale, venice biennale, user biennale, biennalearte2019, biennale arte, art gardens, biennale httpurl, biennial contemporary, biennale venezia, biennalearte2019 user
- [TOP] other sites 1: beach, venice beach, beach httpurl, beach boardwalk, venicebeach california, california sunset, caminomasqueunloco losangeles, beach bordwalk, sunset venice, los

Appendix C. Nomenclature

Tables C.1 gives an overview of the mathematical notations used in this paper.

Table C.1
The nomenclature of mathematical notations used in alphabetic order.

Symbol	Data type/Shape	Description
$\mathcal{A}, \mathcal{B}$	Sets of objects	Generic sets.
$\boldsymbol{c}$	Vector of non-negative integers $\boldsymbol{c} := [c_j]_{\|\mathcal{C}\|\times 1} \in \mathbb{N}^{\|\mathcal{C}\|\times 1}, c_j = \|\{\mathfrak{d}_i \| \mathfrak{c}_i = \zeta_j\}\|$	The number of tweets that are posted in the cities from the set $\mathcal{C}$.
$\boldsymbol{c}_\mathrm{B}, \boldsymbol{c}_\mathrm{D}, \boldsymbol{c}_\mathrm{A}$	Vectors of non-negative integers $\boldsymbol{c}_\mathrm{B}, \boldsymbol{c}_\mathrm{D}, \boldsymbol{c}_\mathrm{A} \in \mathbb{N}^{\|\mathcal{C}\|\times 1}$, $\boldsymbol{c}_\mathrm{B} + \boldsymbol{c}_\mathrm{D} + \boldsymbol{c}_\mathrm{A} = \boldsymbol{c}$	The number of tweets that are posted in each city before, during, and after the event.
$\mathcal{C}$	A set of objects $\mathcal{C} = \{\zeta_0, \zeta_1, \ldots, \zeta_{\|\mathcal{C}\|-1}\}$	The unique names of the cities in the dataset.
$\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2$	Sets of objects $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2 \subset \mathcal{C}, \mathcal{C}_0 = \{\zeta_0\}$	The cities that are the ones where the events happened ($\mathcal{C}_0$), from the same country ($\mathcal{C}_1$), or from far beyond ($\mathcal{C}_2$).
χ^2	Scalar value	The Chi-square statistics of two distributions.
$\boldsymbol{d}$	Vector of non-negative floats $\boldsymbol{d} := [d_i]_{K\times 1} \in \mathbb{R}^{K\times 1}$	The geodesic distances of the cities to the city where the event happened (ζ_0).
df	Scalar value	The degree of freedom.
$df*$	Scalar value	The minimum of the number of rows or columns minus 1 for a two-level Chi Square test.
$\mathfrak{d}_i, \mathfrak{D}$	Object tuples $\mathfrak{d}_i = (\mathcal{S}_i, \mathcal{O}_i, \mathfrak{u}_i, \mathfrak{t}_i, \mathfrak{l}_i)$, $\mathfrak{d}_i \in \mathfrak{D} = \{\mathfrak{d}_0, \mathfrak{d}_1, \ldots, \mathfrak{d}_{K-1}\}$	The tuple of all raw data (sentences, ID of other associated tweets, user ID, timestamp, and geo-location) from one sample point.
$\mathcal{E}^\mathrm{CONV}, \mathcal{E}^\mathrm{USER}$	Sets of tuples $\mathcal{E} = \mathcal{E}^\mathrm{CONV} \bigcup \mathcal{E}^\mathrm{USER}$	The link sets denoting respectively the conversational links and the user links among the tweets.
$\mathcal{G}^\mathrm{MULT}$	Directed multi-graph $\mathcal{G} = (\mathcal{V}, \{\mathcal{E}^\mathrm{CONV}, \mathcal{E}^\mathrm{USER}\})$	The graph including the conversational links and user links among the tweets.
$g_\mathrm{BERT}, g_\mathrm{ULMFiT}$	Models as end-to-end functions returning logit vectors	The pre-trained BERT and ULMFiT models on WHOSe Heritage datasets.
H	Scalar value	The statistics of the Kruskal–Wallis H tests.
i, i'	Integer indices $i, i' \in \{0, 1, 2, \ldots, K-1\} \subset \mathbb{N}$	The index of samples in the dataset $\mathfrak{D}$ of one case heritage-related event.
j	Integer indices $j \in \{0, 1, 2, \ldots, \|\mathcal{C}\|-1\} \subset \mathbb{N}$	The index of cities in the set $\mathcal{C}$ of all unique names of the cities.
k	Integer indices $k \in \{0, 1, 2, \ldots, \|\mathcal{T}\|-1\} \subset \mathbb{N}$	The index of timestamps in the ordered set $\mathcal{T}$ of all unique hours from one case city.
$\boldsymbol{K}^\mathrm{OUV}$	Matrix of floats $\boldsymbol{K}^\mathrm{OUV} = [\boldsymbol{\kappa}_i^\mathrm{OUV}]_{2\times\|\mathcal{V}\|}$	The confidence indicator matrix for OUV labels including the top-n confidence and agreement between BERT and ULMFiT models.
$\boldsymbol{k}^\mathrm{EMS}$	Vector of Boolean's $\boldsymbol{k}^\mathrm{EMS} = [\kappa_i^\mathrm{EMS}]_{1\times\|\mathcal{V}\|}$, $\kappa_i^\mathrm{EMS} = \kappa_i^\mathrm{EM} \wedge \kappa_i^\mathrm{SE}, \kappa_i^\mathrm{EMS}, \kappa_i^\mathrm{EM}, \kappa_i^\mathrm{SE} \in \{0, 1\}$	The confidence indicator vector of emotion labels that shows both a consistent emotion prediction ($\kappa_i^\mathrm{EM} = 1$) and a similar sentiment prediction ($\kappa_i^\mathrm{SE} = 1$) with different models.
K	Integer $K = \|\mathfrak{D}\|$	The sample size (number of posts) collected in one case event.
$\boldsymbol{l}$	Vector of floats	A generic vector.

(*continued on next page*)

Table C.1 (*continued*).

$\mathfrak{l}_i$	Tuple of floats $\mathfrak{l}_i = (\mathfrak{x}_i, \mathfrak{y}_i, \mathfrak{c}_i)$	The geographical coordinate of latitude ($\mathfrak{y}_i$) and longitude ($\mathfrak{x}_i$) and city name ($\mathfrak{c}_i$) as location of one sample.
m	Integer indices $m \in \{0, 1, 2, \ldots, \|\mathcal{Z}\| - 1\} \subset \mathbb{N}$	The index of generated topics $\mathcal{Z}$ from topic modelling.
$\max(\boldsymbol{l}, n)$	Function returning a float with a vector and an integer as inputs	The function returning the value of the n_{th} largest element of a vector $\boldsymbol{l}$.
n	Scalar value	The sample size in a statistical test.
$\boldsymbol{n}$	Vector of integers $\boldsymbol{n} = [1, 2, 3..., \|C\|]^{\mathrm{T}}$	The ranking vector of the ordered set C.
$\mathcal{O}_i$	A set of tuples or an empty set $\mathcal{O}_i = \{\mathfrak{d}_{i'} \| \mathfrak{d}_{i'} \in \mathcal{D}\}$ or $\mathcal{O}_i = \emptyset$	All the tweets that are associated with the tweet $\mathfrak{d}_i$.
$\mathrm{OUV}(\mathcal{A}, \mathcal{B})$	A function returning a scalar with sets as inputs	The function calculating the Intersection over Union of two sets.
p	Scalar value	The significance of a statistical test.
$\mathcal{S}_i$	Set of strings $\mathcal{S}_i = \{s_i^{(1)}, s_i^{(2)}, \ldots, s_i^{(\|\mathcal{S}_i\|)}\}$	The processed textual tweet data as a set of individual sentences that have been translated into English.
$\mathcal{T}$	An ordered set $\mathcal{T} = \{\tau_0, \tau_1, \ldots, \tau_{\|\mathcal{T}\|-1}\}$	The ordered set of all unique timestamps from one case event.
$\mathcal{T}_{\mathrm{B}}, \mathcal{T}_{\mathrm{D}}, \mathcal{T}_{\mathrm{A}}$	Ordered subsets $\mathcal{T}_{\mathrm{B}}, \mathcal{T}_{\mathrm{D}}, \mathcal{T}_{\mathrm{A}} \subset \mathcal{T}$	The ordered set of all unique timestamps before, during, and after the event.
$\text{top-n}(\boldsymbol{l}, n)$	Function returning a set with a vector and an integer as inputs	The function returning the index set of the largest n elements in the vector $\boldsymbol{l}$.
τ_k	Timestamp $\tau_k \in \mathcal{T}$	A timestamp in the ordered set $\mathcal{T}$ of all unique timestamps.
$\mathfrak{t}_i$	Timestamp $\mathfrak{t}_i \in \mathcal{T}$	A timestamp indexed with sample ID in the ordered set $\mathcal{T}$ of all unique timestamps.
$\boldsymbol{t}$	Vector of non-negative integers $\boldsymbol{t} := [t_k]_{\|\mathcal{T}\|\times 1} \in \mathbb{N}^{\|\mathcal{T}\|\times 1}, t_k = \|\{\mathfrak{d}_i \| \mathfrak{t}_i = \tau_k\}\|$	The number of tweets that are posted at each unique timestamp.
$\boldsymbol{\Theta}_{\mathrm{BERT}}, \boldsymbol{\Theta}_{\mathrm{ULMFiT}}$	Parameters	Model parameters for the BERT and ULMFiT models.
$\mathcal{U}$	An ordered set $\mathcal{U} = \{\mu_0, \mu_1, \ldots, \mu_{\|\mathcal{U}\|-1}\}$	The ordered set of all unique users from one case event.
$\mathfrak{u}_i$	User ID object $\mathfrak{u}_i \in \mathcal{U}$	An instance of user indexed with sample ID in the ordered set $\mathcal{U}$ of all unique users.
U	Scalar value	The statistics of the Mann–Whitney U tests.
V	Scalar value	The Cramer's V as effect size for Chi Square tests.
$\mathcal{V}$	A set of nodes $\mathfrak{d}_i \in \mathcal{V}, \mathcal{V} \subset \mathcal{D}$	The set of all nodes of tweets in a case event that are not isolated.
$\mathcal{V}^{\mathrm{OUV}}, \mathcal{V}^{\mathrm{EMS}}, \mathcal{V}^{\mathrm{TOP}}$	Sets of nodes $\mathcal{V}^{\mathrm{OUV}}, \mathcal{V}^{\mathrm{EMS}}, \mathcal{V}^{\mathrm{TOP}} \subset \mathcal{V} \subset \mathcal{D}$	The sets of filtered tweets that are found to give valid predictions on OUV, emotion, and topic labels.
$(\mathfrak{x}_i, \mathfrak{y}_i)$	Geographical coordinates	The latitude and longitude of the tweet $\mathfrak{d}_i$.
(x_0, y_0)	Geographical coordinates	The latitude and longitude of the city ζ_0 where the event happened.
$\boldsymbol{Y}^{\mathrm{OUV}}$	Matrix of floats $\boldsymbol{Y}^{\mathrm{OUV}} = [\boldsymbol{y}_i^{\mathrm{OUV}}]_{11\times K}$	The OUV labels of tweets as probability distributions on 10 OUV selection criteria and an additional negative class, as the average of prediction from BERT and ULMFiT models.
$\boldsymbol{y}_i^{\mathrm{BERT}}, \boldsymbol{y}_i^{\mathrm{UMLFiT}}$	Logit vector of floats $\boldsymbol{y}_i^{\mathrm{BERT}} \in [0,1]^{11\times 1}, \boldsymbol{y}_i^{\mathrm{ULMFiT}} \in [0,1]^{11\times 1}$	Predicted OUV labels for the tweet $\mathfrak{d}_i$ by BERT and ULMFiT models
$\boldsymbol{y}_i^{\mathrm{EM}(0)}, \boldsymbol{y}_i^{\mathrm{EM}(1)}$	Logit vector of Floats $\boldsymbol{y}_i^{\mathrm{EM}(0)} \in [0,1]^{7\times 1}, \boldsymbol{y}_i^{\mathrm{EM}(1)} \in [0,1]^{6\times 1}$	Predicted emotion labels for the tweet $\mathfrak{d}_i$ by pysentimiento and BERTweet emotion models.
$\boldsymbol{y}_i^{\mathrm{SE}(0)}, \boldsymbol{y}_i^{\mathrm{SE}(1)}$	Logit vector of Floats $\boldsymbol{y}_i^{\mathrm{SE}(0)} \in [0,1]^{3\times 1}, \boldsymbol{y}_i^{\mathrm{SE}(1)} \in [0,1]^{3\times 1}$	Predicted sentiment labels for the tweet $\mathfrak{d}_i$ by pysentimiento and BERTweet sentiment models.
$\boldsymbol{y}_i^{\mathrm{TOP}}$	Logit vector of floats $\boldsymbol{y}_i^{\mathrm{TOP}} = \left[y_{i,m}^{\mathrm{TOP}}\right]_{\|\mathcal{Z}\|\times 1} \in [0,1]^{\|\mathcal{Z}\|\times 1}$	Predicted topic labels for the tweet $\mathfrak{d}_i$ with topic modelling from BERTopic.
$\mathcal{Y}^{\mathrm{EMS}}$	Array of sets $\mathcal{Y}^{\mathrm{EMS}} = \left[\mathcal{Y}_i^{\mathrm{EMS}}\right]$	The array of final emotion labels for all the tweets in $\mathcal{V}$, containing the top-1 emotions and top-1 sentiments if the prediction is valid, otherwise empty.
$\mathcal{Y}^{\mathrm{OUV}}$	Array of sets $\mathcal{Y}^{\mathrm{OUV}} = \left[\mathcal{Y}_i^{\mathrm{OUV}}\right]$	The array of final OUV labels for all the tweets in $\mathcal{V}$, containing the top-3 OUV selection criteria if the prediction is valid, otherwise empty.

(*continued on next page*)

Table C.1 (*continued*).

$\mathcal{Y}^{\text{TOP}}$	Array of sets $\mathcal{Y}^{\text{TOP}} = [\mathcal{Y}_i^{\text{TOP}}]$	The array of final topic labels for all the tweets in $\mathcal{V}$, containing the topic name that is within key topics $\mathcal{Z}^s$ and has a higher probability than 0.5, otherwise empty.
$\mathcal{Z}$	A set of objects $\mathcal{Z} = \{z_m \mid m = -1, 0, 1, \ldots, \lvert\mathcal{Z}\rvert - 2\}$	The set of the generated topics obtained with BERTopic topic modelling, where z_{-1} is the "noise" topic.
$\mathcal{Z}^s$	A subset of objects $\mathcal{Z}^s \in \mathcal{Z}$	A subset of the generated key topics obtained with BERTopic topic modelling that are interesting and informative for heritage management.
ζ_0	An object $\zeta_0 \in \mathcal{C}, \mathcal{C}_0 = \{\zeta_0\}$	The name of the city where the event happened.
ζ_j	An object $\zeta_j \in \mathcal{C}$	The name of a city that is one instance of the set $\mathcal{C}$.

References

[1] V. Monteiro, R. Henriques, M. Painho, E. Vaz, Sensing world heritage an exploratory study of Twitter as a tool for assessing reputation, in: B. Murgante, S. Misra, A.M.A.C. Rocha, C. Torre, J.G. Rocha, M.I. Falcao, D. Taniar, B.O. Apduhan, O. Gervasi (Eds.), COMPUTATIONAL SCIENCE and ITS APPLICATIONS - ICCSA 2014, PT II, Univ Minho; Univ Perugia; Univ Basilicata; Monash Univ; Kyushu Sangyo Univ; Assoc Portuguesa Investigacao Operac, 2014, pp. 404–419.

[2] V. Monteiro, M. Painho, E. Vaz, Is the heritage really important? a theoretical framework for heritage reputation using citizen sensing, Habitat Int. 45 (2015) 156–162.

[3] A. Chianese, F. Marulli, F. Piccialli, Cultural heritage and social pulse: A semantic approach for CH sensitivity discovery in social media data, in: Proceedings - 2016 IEEE 10th International Conference on Semantic Computing, ICSC 2016, Institute of Electrical and Electronics Engineers Inc., 2016, pp. 459–464, http://dx.doi.org/10.1109/ICSC.2016.50.

[4] N. Bai, P. Nourian, A. Pereira Roders, Global citizens and world heritage: Social inclusion of online communities in heritage planning, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLVI-M-1-2021 (2021) 23–30, http://dx.doi.org/10.5194/isprs-archives-XLVI-M-1-2021-23-2021.

[5] L. Li, M. Bensi, G. Baecher, Exploring the potential of social media crowdsourcing for post-earthquake damage assessment, Int. J. Disaster Risk Reduct. 98 (2023) 104062.

[6] H. Stovel, Risk Preparedness: A Management Manual for World Cultural Heritage, Technical Report, ICCROM, in collaboration with ICOMOS, UNESCO, and the World Heritage Center, Rome, 1998.

[7] J. Jokilehto, Iccrom's involvement in risk preparedness, J. Am. Inst. Conserv. 39 (2000) 173–179.

[8] UNESCO, Strategy for Reducing Risks At World Heritage Properties, Technical Report, UNESCO World Heritage Centre, 2007.

[9] UNESCO, Managing Disaster Risks for World Heritage, Technical Report, UNESCO World Heritage Centre, 2010.

[10] UNESCO, Recommendation on the Historic Urban Landscape, Technical Report, UNESCO, Paris, 2011.

[11] ICCROM, A Guide To Risk Management of Cultural Heritage, Technical Report, ICCROM, 2016.

[12] ICOMOS-ICCROM, Guidance on Post-Disaster and Post-Conflict Recovery and Reconstruction for Heritage Places of Cultural Significance and World Heritage Cultural Properties, Technical Report, ICOMOS-ICCROM, 2024.

[13] J.A. Tucker, A. Guess, P. Barberá, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, B. Nyhan, Social media, political polarization, and political disinformation: A review of the scientific literature, in: Political Polarization, and Political Disinformation: A Review of the Scientific Literature (March 19 2018), 2018.

[14] T.M. Stevens, N. Aarts, A. Dewulf, Using emotions to frame issues and identities in conflict: farmer movements on social media, Negot. Confl. Manag. Res. (2020).

[15] X. Zhai, Q. Luo, L. Wang, Why tourists engage in online collective actions in times of crisis: Exploring the role of group relative deprivation, J. Destination Mark. Manag. 16 (2020) http://dx.doi.org/10.1016/j.jdmm.2020.100414.

[16] A. Schroeder, L. Pennington-Gray, H. Donohoe, S. Kiousis, Using social media in times of crisis, J. Travel Tour. Mark. 30 (2013) 126–143, http://dx.doi.org/10.1080/10548408.2013.751271.

[17] C. Garduño Freeman, J.A. Gonzalez Zarandona, Digital spectres: the notre-dame effect, Int. J. Herit. Stud. 27 (2021) 1264–1277.

[18] C. Lipizzi, L. Iandoli, J.E.R. Marquez, Extracting and evaluating conversational patterns in social media: A socio-semantic analysis of customers' reactions to the launch of new products using twitter streams, Int. J. Inf. Manage. 35 (2015) 490–503.

[19] Y. Zhang, T. Cheng, Graph deep learning model for network-based predictive hotspot mapping of sparse spatio-temporal events, Comput. Environ. Urban Syst. 79 (2020) 101403.

[20] N. Bai, P. Nourian, R. Luo, A. Pereira Roders, Heri-graphs: A dataset creation framework for multi-modal machine learning on graphs of heritage values and attributes with social media, ISPRS Int. J. Geo-Inf. 11 (2022) http://dx.doi.org/10.3390/ijgi11090469.

[21] M. Foroughi, B. de Andrade, A.P. Roders, Capturing public voices: The role of social media in heritage management, Habitat Int. 142 (2023) 102934.

[22] UNESCO, Implementing New Visions: A Guidebook for Action on the List of World Heritage in Danger, Technical Report, UNESCO World Heritage Centre, 2024.

[23] A. Pereira Roders, The historic urban landscape approach in action: Eight years later, in: Reshaping Urban Conservation, Springer, 2019, pp. 21–54, http://dx.doi.org/10.1007/978-981-10-8887-2_2.

[24] T. Cheng, T. Wicks, Event detection using twitter: A spatio-temporal approach, PLoS One 9 (2014) e97807.

[25] Z. Shi, L.S. Pun-Cheng, Spatiotemporal data clustering: a survey of methods, ISPRS Int. J. Geo-Inf. 8 (2019) 112.

[26] J.O. Arjona, Analysis of the space-temporal patterns of events from twitter data: The case of madrid 2017 world pride, Estudios Geograficos 81 (2020) http://dx.doi.org/10.3989/estgeogr.202047.027.

[27] J. Kersten, F. Klan, What happens where during disasters? a workflow for the multifaceted characterization of crisis events based on twitter data, J. Conting. Crisis Manag. 28 (2020) 262–280.

[28] Y. George, S. Karunasekera, A. Harwood, K.H. Lim, Real-time spatio-temporal event detection on geotagged social media, J. Big Data 8 (2021) 91.

[29] I. Afyouni, A. Khan, Z.A. Aghbari, Deep-eware: spatio-temporal social event detection using a hybrid learning model, J. Big Data 9 (2022) http://dx.doi.org/10.1186/s40537-022-00636-w.

[30] P. Cucco, G. Di Ruocco, L. La Rana, Proposal of an innovative model for fire prevention assessment in cultural heritage protection. research study in italy, Int. J. Disaster Risk Reduct. 97 (2023) 104066.

[31] M. Meghraoui, R. Sbeinati, Archeoseismology and the lost villages in northern syria, the impact of large earthquakes on cultural heritage, in: Sustainable Conservation of UNESCO and Other Heritage Sites Through Proactive Geosciences, Springer, 2023, pp. 445–461.

[32] J. Sofaer, B. Davenport, M.L.S. Sørensen, E. Gallou, D. Uzzell, Heritage sites, value and wellbeing: learning from the covid-19 pandemic in england, Int. J. Herit. Stud. 27 (2021) 1117–1132.
[33] M. Ginzarly, F.J. Srour, Cultural heritage through the lens of covid-19, Poetics 92 (2022) 101622.
[34] M. Naramski, A.R. Szromek, K. Herman, G. Polok, Assessment of the activities of european cultural heritage tourism sites during the covid-19 pandemic, J. Open Innov.: Technol. Mark. Complexity 8 (2022) 55.
[35] M. Tenzer, Tweets in the peak: Twitter analysis-the impact of covid-19 on cultural landscapes, Int. Archaeology 59 (2022).
[36] P. Kumar, Learning from the Past and Preparing for the Future: Cases and Tools for Cultural Heritage During Disasters (Ph.D. thesis), IMT School for Advanced Studies Lucca, 2019.
[37] P. Kumar, Crowdsourcing to rescue cultural heritage during disasters: A case study of the 1966 florence flood, Int. J. Disaster Risk Reduct. 43 (2020) 101371.
[38] P. Kumar, F. Ofli, M. Imran, C. Castillo, Detection of disaster-affected cultural heritage sites from social media images using deep learning techniques, J. Comput. Cult. Herit. (JOCCH) 13 (2020) 1–31.
[39] P. Kumar, Twitter, disasters and cultural heritage: A case study of the 2015 nepal earthquake, J. Conting. Crisis Manag. 28 (2020) 453–465.
[40] Y. Huang, Y. Li, J. Shan, Spatial–temporal event detection from geo-tagged tweets, ISPRS Int. J. Geo-Inf. 7 (2018) 150.
[41] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.
[42] Z. Ghaemi, M. Farnaghi, A varied density-based clustering approach for event detection from heterogeneous twitter data, ISPRS Int. J. Geo-Inf. 8 (2019) 82.
[43] M. Farnaghi, Z. Ghaemi, A. Mansourian, Dynamic spatio-temporal tweet mining for event detection: a case study of hurricane florence, Int. J. Disaster Risk Sci. 11 (2020) 378–393.
[44] M. Rani, S. Kaushal, Geoclust: Feature engineering based framework for location-sensitive disaster event detection using ahp-topsis, Expert Syst. Appl. 210 (2022) http://dx.doi.org/10.1016/j.eswa.2022.118461.
[45] M. Fedoryszak, B. Frederick, V. Rajaram, C. Zhong, Real-time event detection on social data streams, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 2774–2782.
[46] O. Yanenko, A. der Weberei, Introducing social distance to st-dbscan, in: Proceedings of the 22nd AGILE Conference, 2019, pp. 1–5.
[47] C. Choi, S.Y. Hong, Mdst-dbscan: A density-based clustering method for multidimensional spatiotemporal data, ISPRS Int. J. Geo-Inf. 10 (2021) 391.
[48] B. Martínez-López, A. Perez, J. Sánchez-Vizcaíno, Combined application of social network and cluster detection analyses for temporal-spatial characterization of animal movements in salamanca, spain, Prev. Vet. Med. 91 (2009) 29–38.
[49] M.A. Costa, M. Kulldorff, Maximum linkage space–time permutation scan statistics for disease outbreak detection, Int. J. Health Geogr. 13 (2014) 1–14.
[50] T.C. Wang, F.K.H. Phoa, A scanning method for detecting clustering pattern of both attribute and structure in social networks, Phys. A 445 (2016) 295–309.
[51] M. Adepeju, Modelling of Sparse Spatio-Temporal Point Process (STPP): An Application in Predictive Policing (Ph.D. thesis), UCL (University College London), 2017.
[52] J. Shen, Profiling and Grouping Space–Time Activity Patterns of Urban Individuals (Ph.D. thesis), UCL (University College London), 2018.
[53] R. Padilha, F.A. Andaló, B. Lavi, L.A. Pereira, A. Rocha, Temporally sorting images from real-world events, Pattern Recognit. Lett. 147 (2021) 212–219.
[54] R. Padilha, F.A. Andaló, L.A. Pereira, A. Rocha, Unraveling the notre-dame cathedral fire in space and time: an x-coherence approach, in: Crime Science and Digital Forensics, CRC Press, 2021, pp. 3–19.
[55] L.C. Passaro, A. Bondielli, P. Dell'Oglio, A. Lenci, F. Marcelloni, In-context annotation of topic-oriented datasets of fake news: A case study on the notre-dame fire event, Inform. Sci. 615 (2022) 657–677.
[56] B.M. Lupo, Patrimônio cultural e cat´ ástrofe: Os concursos internacionais não-oficiais realizados para a notre dame de paris após o incêndio de 2019, Herança-Rev. História, Património Cult. 4 (2021) 018–038.
[57] S.O. Molina, M.Á.O. Molina, Notre-dame de parís. dov'era e com'era: la réplica que habita en la ruina, Loggia, Arquitectura Restauración 34 (2021) 8–27.
[58] D. Card, A. Boydstun, J.H. Gross, P. Resnik, N.A. Smith, The media frames corpus: Annotations of frames across issues, in: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), 2015, pp. 438–444.
[59] S. Roy, D. Goldwasser, Weakly supervised learning of nuanced frames for analyzing polarization in news media, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2020, pp. 7698–7716.
[60] C.W. van Eck, B.C. Mulder, A. Dewulf, Online climate change polarization: Interactional framing analysis of climate change blog comments, Sci. Commun. 42 (2020) 454–480.
[61] Z. Wang, X. Ye, Social media analytics for natural disaster management, Int. J. Geogr. Inf. Sci. 32 (2018) 49–72.
[62] Y. Feng, X. Huang, M. Sester, Extraction and analysis of natural disaster-related vgi from social media: review, opportunities and challenges, Int. J. Geogr. Inf. Sci. 36 (2022) 1275–1316.
[63] A. Bruns, J.E. Burgess, K. Crawford, F. Shaw, # Qldfloods and QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods, Technical Report, Cci. ARC Centre of Excellence for Creative Industries and Innovation, 2012.
[64] Y. Kryvasheyeu, H. Chen, E. Moro, P. Van Hentenryck, M. Cebrian, Performance of social network sensors during hurricane sandy, PLoS One 10 (2015) e0117288.
[65] P. Kumar, Social media, disasters, and cultural heritage: An analysis of twitter images of the 2015 nepal earthquake, Vis. Commun. Q. 29 (2022) 34–46.
[66] A.B. Khajwal, A. Noshadravan, An uncertainty-aware framework for reliable disaster damage assessment via crowdsourcing, Int. J. Disaster Risk Reduct. 55 (2021) 102110.
[67] X. Romao, C. Bertolin, Risk protection for cultural heritage and historic centres: Current knowledge and further research needs, Int. J. Disaster Risk Reduct. 67 (2022) 102652.
[68] Y. Zhang, R. Zong, Z. Kou, L. Shang, D. Wang, Collablearn: An uncertainty-aware crowd-ai collaboration system for cultural heritage damage assessment, IEEE Trans. Comput. Soc. Syst. 9 (2021) 1515–1529.
[69] V. Lorini, P. Rufolo, C. Castillo, Venice was flooding. one tweet at a time, Proc. ACM Hum.-Comput. Interact. 6 (2022) 1–16.
[70] N. Demir, Ç. Çlingiroğlu, Assessing the impact of the türkiye february 2023 earthquakes on cultural heritage sites: A multi-disciplinary approach utilizing aria maps and social media collaboration, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 48 (2024) 117–122.
[71] Y. Praticò, J. Ochsendorf, S. Holzer, R.J. Flatt, Post-fire restoration of historic buildings and implications for notre-dame de paris, Nature Mater. 19 (2020) 817–820.
[72] C. Ferrarin, M. Bajo, A. Benetazzo, L. Cavaleri, J. Chiggiato, S. Davison, S. Davolio, P. Lionello, M. Orlić, G. Umgiesser, Local and large-scale controls of the exceptional venice floods of november 2019, Prog. Oceanogr. 197 (2021) 102628.
[73] O. Kounadi, T.J. Lampoltshammer, M. Leitner, T. Heistracher, Accuracy and privacy aspects in free online reverse geocoding services, Cartogr. Geogr. Inf. Sci. 40 (2013) 140–153.
[74] K. Clemens, Geocoding with openstreetmap data, GEOProcessing 2015 (2015) 10.
[75] S. Bird, E. Klein, E. Loper, Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, O'Reilly Media, Inc, 2009.

[76] J. Devlin, M.W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171–4186, http://dx.doi.org/10.18653/v1/N19-1423, URL: https://aclanthology.org/N19-1423.
[77] M. Cristelli, M. Batty, L. Pietronero, There is more than a power law in zipf, Sci. Rep. 2 (2012) 1–7.
[78] I. Moreno-Sánchez, F. Font-Clos, Á. Corral, Large-scale analysis of zipf's law in english texts, PLoS One 11 (2016) e0147073.
[79] J. Jokilehto, What Is OUV? Defining the Outstanding Universal Value of Cultural World Heritage Properties, Technical Report, ICOMOS, ICOMOS Berlin, 2008.
[80] UNESCO, Operational Guidelines for the Implementation of the World Heritage Convention, Technical Report July, UNESCO World Heritage Centre, 2008.
[81] IUCN, ICOMOS, ICROM, WHC, Guidance on the Preparation of Retrospective Statements of Outstanding Universal Value for World Heritage Properties, Technical Report, IUCN and ICOMOS and ICROM and WHC, 2010.
[82] A. Pereira Roders, Re-Architecture: Lifespan Rehabilitation of Built Heritage (Ph.D. thesis), Technische Universiteit Eindhoven, 2007, http://dx.doi.org/10.6100/IR751759.
[83] A. Tarrafa Silva, A. Pereira Roders, The cultural significance of World Heritage cities : Portugal as case study, in: Heritage and Sustainable Development, Évora, Portugal, 2010, pp. 255–263, http://dx.doi.org/10.13140/2.1.1152.0800.
[84] L. Veldpaus, Historic Urban Landscapes: Framing the Integration of Urban and Heritage Planning in Multilevel Governance (Ph.D. thesis), Technische Universiteit Eindhoven, 2015.
[85] J. Howard, S. Ruder, Universal language model fine-tuning for text classification, in: I. Gurevych, Y. Miyao (Eds.), Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July (2018) 15-20, Volume 1: Long Papers, Association for Computational Linguistics, 2018, pp. 328–339, http://dx.doi.org/10.18653/v1/P18-1031.
[86] N. Bai, R. Luo, P. Nourian, A. Pereira Roders, WHOSe heritage: Classification of UNESCO world heritage statements of outstanding universal value with soft labels, in: Findings of the Association for Computational Linguistics: EMNLP 2021, Association for Computational Linguistics, Punta Cana, Dominican Republic, 2021, pp. 366–384, http://dx.doi.org/10.18653/v1/2021.findings-emnlp.34.
[87] F.A. Acheampong, C. Wenyu, H. Nunoo-Mensah, Text-based emotion detection: Advances, challenges, and opportunities, Eng. Rep. 2 (2020) e12189.
[88] J. Eisenstein, Natural Language Processing, MIT Press, 2018.
[89] D. Rao, B. McMahan, Natural Language Processing with PyTorch - Build Intelligent Language Applications using Deep Learning, O'Reilly Media, Inc., 2019, arXiv:arXiv:1011.1669v3.
[90] D. Jurafsky, J.H. Martin, Speech and language processing an intoduction to natural language processing, computational liguistics, and speech recognition, 2020, http://dx.doi.org/10.4324/9780203461891_chapter_3, (Third Edition Draft).
[91] A. Dickinger, L. Lalicic, An analysis of destination brand personality and emotions: a comparison study, Inf. Technol. Tour. 15 (2016) 317–340.
[92] M. Mazloom, B. Hendriks, M. Worring, Multimodal context-aware recommender for post popularity prediction in social media, in: Proceedings of the on Thematic Workshops of ACM Multimedia 2017, 2017, pp. 236–244.
[93] A. Nenko, M. Petrova, Emotional geography of st. petersburg: detecting emotional perception of the city space, in: Digital Transformation and Global Society: Third International Conference, DTGS 2018, St. Petersburg, Russia, May 30–June 2 2018, Revised Selected Papers, Part II 3, Springer, 2018, pp. 95–110.
[94] M. Afzaal, M. Usman, A.C. Fong, S. Fong, Multiaspect-based opinion classification model for tourist reviews, Expert Syst. 36 (2019) e12371.
[95] J. Pan, N. Mou, W. Liu, Emotion analysis of tourists based on domain ontology, in: Proceedings of the 2019 International Conference on Data Mining and Machine Learning, 2019, pp. 146–150.
[96] V. Taecharungroj, B. Mathayomchan, Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand, Tour. Manag. 75 (2019) 550–568, http://dx.doi.org/10.1016/j.tourman.2019.06.020.
[97] P. Ekman, An argument for basic emotions, Cogn. Emot. 6 (1992) 169–200.
[98] Z.H. Zhou, Ensemble Methods: Foundations and Algorithms, CRC Press, 2012.
[99] M. García-Vega, M. Díaz-Galiano, M. García-Cumbreras, F. Del Arco, A. Montejo-Ráez, S. Jiménez-Zafra, E. Martínez Cámara, C. Aguilar, M. Cabezudo, L. Chiruzzo, et al., Overview of tass 2020: Introducing emotion detection, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2020) Co-Located with 36th Conference of the Spanish Society for Natural Language Processing (SEPLN 2020), Málaga, Spain, 2020, pp. 163–170.
[100] J.M. Pérez, J.C. Giudici, F. Luque, Pysentimiento: A python toolkit for sentiment analysis and socialnlp tasks, 2021, arXiv:2106.09462.
[101] J.M. Pérez, D.A. Furman, L. Alonso Alemany, F.M. Luque, RoBERTuito: a pre-trained language model for social media text in Spanish, in: Proceedings of the Thirteenth Language Resources and Evaluation Conference, European Language Resources Association, Marseille, France, 2022, pp. 7235–7243, URL: https://aclanthology.org/2022.lrec-1.785.
[102] D.Q. Nguyen, T. Vu, A.T. Nguyen, Bertweet: A pre-trained language model for english tweets, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 2020, pp. 9–14.
[103] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T.L. Scao, S. Gugger, M. Drame, Q. Lhoest, A.M. Rush, Transformers: State-of-the-art natural language processing, in: Q. Liu, D. Schlangen (Eds.), Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, EMNLP 2020 - Demos, Online, November 16-20, 2020, 2020, pp. 38–45, http://dx.doi.org/10.18653/v1/2020.emnlp-demos.6.
[104] M. Grootendorst, Bertopic: Neural topic modeling with a class-based tf-idf procedure, 2022, arXiv preprint arXiv:2203.05794.
[105] L. McInnes, J. Healy, S. Astels, Hdbscan: Hierarchical density based clustering, J. Open Source Softw. 2 (2017) 205.
[106] F.J. Gravetter, L.B. Wallnau, L.A.B. Forzano, J.E. Witnauer, Essentials of Statistics for the Behavioral Sciences, Cengage Learning, 2020.
[107] Council of Europe, Convention on the Value of Cultural Heritage for Society (Faro Convention), Technical Report, Council of Europe, Faro, 2005.
[108] E. Bozdag, Q. Gao, G.J. Houben, M. Warnier, Does offline political segregation affect the filter bubble? an empirical analysis of information diversity for dutch and turkish twitter users, Comput. Hum. Behav. 41 (2014) 405–415.
[109] M. Suzuki, Y. Yamamoto, Analysis of relationship between confirmation bias and web search behavior, in: Proceedings of the 22nd International Conference on Information Integration and Web-Based Applications & Services, 2020, pp. 184–191.
[110] L. Meskell, States of conservation: protection, politics, and pacting within unesco's world heritage committee, Anthropol. Q. (2014) 217–243.
[111] Musteaţă S., The state of conservation and periodic reporting–a way for better preservation and sustainable development of the world heritage sites. PLURAL. History. Culture. Society, J. History Geogr. Dep. Ion Creangă State Pedagog. Univ. 8 (2020) 227–242.
[112] M. Won, P. Murrieta-Flores, B. Martins, Ensemble named entity recognition (ner): evaluating ner tools in the identification of place names in historical corpora, Front. Digit. Humanit. 5 (2018) 2.
[113] W. Zhang, J. Gelernter, Geocoding location expressions in twitter messages: A preference learning method, J. Spatial Inf. Sci. 9 (2014) 37–70.
[114] B. Dahal, S.A. Kumar, Z. Li, Topic modeling and sentiment analysis of global climate change tweets, Soc. Netw. Anal. Min. 9 (2019) 1–20.
[115] Y. Luo, D. Card, D. Jurafsky, Detecting stance in media on global warming, 2020, arXiv preprint arXiv:2010.15149.
[116] A. Dewulf, R. Bouwen, Issue framing in conversations for change: Discursive interaction strategies for doing differences, J. Appl. Behav. Sci. 48 (2012) 168–193.
[117] E. Rojas-Padilla, T. Metze, K. Termeer, Seeing the visual: A literature review on why and how policy scholars would do well to study influential visualizations, Policy Stud. Yearb. 12 (2022) 103–136.

[118] N. Bai, P. Nourian, R. Luo, T. Cheng, A.P. Roders, Screening the stones of venice: Mapping social perceptions of cultural significance through graph-based semi-supervised classification, ISPRS J. Photogramm. Remote Sens. 203 (2023) 135–164, http://dx.doi.org/10.1016/j.isprsjprs.2023.07.018.
[119] B. Andrade, I can see through the waters eyes. covid-19 in heritage cities: Citizen participation and self-organization for greater conservation and sustainability: The case of venezia pulita (clean venice), in: Living (World) Heritage Cities: Opportunities, Challenges, and Future Perspectives of People-Centered Approaches in Dynamic Historic Urban Landscapes, Sidestone Press, 2022, pp. 199–212.
[120] R. Jigyasu, Managing cultural heritage in the face of climate change, J. Int. Aff. 73 (2019) 87–100.
[121] E. Sesana, A.S. Gagnon, A. Bonazza, J.J. Hughes, An integrated approach for assessing the vulnerability of world heritage sites to climate change impacts, J. Cult. Herit. 41 (2020) 211–224.
[122] L.J. Durrant, A.N. Vadher, J. Teller, Disaster risk management and cultural heritage: The perceptions of european world heritage site managers on disaster risk management, Int. J. Disaster Risk Reduct. 89 (2023) 103625.
[123] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP, Association for Computational Linguistics, 2019, pp. 3982–3992.