The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

MAPPING URBAN HERITAGE IMAGES WITH SOCIAL MEDIA DATA AND ARTIFICIAL INTELLIGENCE, A CASE STUDY IN TESTACCIO, ROME

N. Bai[1], M. Ducci[2], R. Mirzikashvili[3], P. Nourian[4], A. Pereira Roders[1]

[1] Delft University of Technology, Delft, the Netherlands - {n.bai, a.r.pereira-roders}@tudelft.nl
[2] Vrije Universiteit Amsterdam, Amsterdam, the Netherlands - m.ducci@vu.nl
[3] Newcastle University, Newcastle upon Tyne, UK - r.mirzikashvili@gmail.com
[4] University of Twente, Enschede, the Netherlands – p.nourian@utwente.nl

KEY WORDS: Historic Urban Landscape, Machine Learning, User-Generated Content, Cultural Heritage, Big Data Analysis

ABSTRACT:

The UNESCO 2011 *Recommendation on the Historic Urban Landscape* promotes to map cultural significance of urban heritage from the perspectives of the general public in pursuit of social inclusion in heritage management. The user-generated information already available on social media platforms in the form of images, comments, and ratings can be considered a rich source for collecting data concerning the tourists' image of destinations and their collective perception of urban cultural heritage. Considering the large amount of unstructured data, artificial intelligence (AI) can construct structured feature vectors therefrom and significantly aid the analysis and collation processes compared to the traditional manual approach for mapping public perception of cultural heritage. This paper presents an exploratory case study conducted in the area of Testaccio, Rome, showcasing the use of AI to map the perceived and narrated urban heritage images using social media data. An image-sharing platform, Flickr, is used to collect thousands of posts containing images and comments in the area, which are further analysed with pre-trained image recognition, natural language processing, and dimensionality reduction algorithms. Results as the urban heritage images are visualised, showing the most significant elements from a public perspective. Such a methodology provides an alternative perspective of viewing the urban heritage attributes as a collection of depicted and posted content. It can contribute as a tool for the documentation of collective attention for inclusive heritage management and local development planning during the designing and policy-making processes.

1. INTRODUCTION

Conventionally, the listing of cultural heritage, especially that of UNESCO World Heritage is determined by authorities and justified by experts, thus mainly a top-down view of the cultural significance (UNESCO, 2008). However, ordinary people including local residents and visiting tourists usually also have their own experiences and opinions on the tangible and intangible aspects that they truly value in a heritage-rich place, providing an alternative bottom-up view (Janssen et al., 2017, Bai et al., 2021a). Such a view as the "image" of a place by the public can be informative for experts during the spatial planning and heritage management decision-making processes, as it adds a potentially more inclusive layer of information concerning the emotional attachment and sense of belonging in a "*lived place*" (Lynch, 1964, Lefebvre, 1991), which might not be directly and/or necessarily recognised as heritage according to the conventional procedure but does positively contribute to the collective memory (Assmann et al., 1995, Bai et al., 2023). Since the adoption of the 2011 *Recommendation on the Historic Urban Landscape* (HUL) by UNESCO, mapping the cultural significance of urban heritage attributes (Veldpaus, 2015) from the perspectives of a broader range of stakeholders including the general public is being recommended (UNESCO, 2011), where tools for knowledge documentation and civic engagement are also actively called for. As previously argued by Bai et al. (2021a), social media platforms and other online digital applications already partially function as a critical resource for constructing such tools to promote inclusive planning, both as an active medium for crowd-sourcing and participatory design (Watkins, 2007, Estellés-Arola and González Ladrón de-Guevara, 2012, Boy and Uitermark, 2017, Ducci et al., 2023), and as a pre-existing database for social sensing, information mapping, and pattern mining (Schich et al., 2014, Ginzarly et al., 2019, Kumar, 2020, Galesic et al., 2021).

The geo-tagged and time-stamped user-generated information already available on social media platforms in the form of images, comments, and ratings is considered an effective source for collecting data concerning the tourist image of destinations (Kang et al., 2021, Cho et al., 2022), the most representative characteristics of urban scenes (Lai et al., 2017, Zhang et al., 2019, Liu et al., 2021), and more specifically, the collective perception of urban cultural heritage (Monteiro et al., 2014, Ginzarly et al., 2019, Bai et al., 2022), all of which could be interpreted as major components of the urban heritage images perceived and expressed by the social media users. However, the amount of user-generated data on social media platforms is usually at a large scale with thousands and even millions of samples that easily exceed the capacity of manual analyses, even if only a general overview is desired.

The current advancements in the field of artificial intelligence (AI), or specifically, the developments of pre-trained machine learning and deep learning models have made semi-automatic analyses of unstructured multi-modal data at scales not only possible but also effective (Deng et al., 2009, LeCun 2015, Vaswani et al., 2017, Baltrušaitis et al., 2018). With large image and language models thoroughly pre-trained on a massive amount of data, structured feature vectors can be easily constructed and/or extracted as effective representations of the raw information, enabling various types of downstream tasks in the application ends at a relatively low cost (Pan et al., 2010, Kang et al., 2021). As a vivid example, ChatGPT already showcased how AI models could interact with different use cases including urban, history, and heritage studies, creating

revolutionary possibilities, even though also raising moral concerns (Batty 2023, Fostikov, 2023, Thorp 2023). Nevertheless, this paper merely considers AI as a toolbox that can significantly aid the analysis and collation processes compared to the traditional manual approach in heritage studies.

This paper will present an exploratory case study conducted in the historic area of Testaccio in Rome, Italy, demonstrating the use of AI to process the collected user-generated social media data, to map the perceived and narrated urban heritage images, and to integrate the results for cultural heritage management and local development planning practices.

2. METHODOLOGY

2.1 Case Study

The Testaccio area is chosen as a case study. Being at the border within the UNESCO World Heritage property "*Historic Centre of Rome, the Properties of the Holy See in that City Enjoying Extraterritorial Rights and San Paolo Fuori le Mura*" (https://whc.unesco.org/en/list/91/, accessed 05 April 2023, under the license CC-BY-SA IGO 3.0), its historical and cultural values are officially justified with Outstanding Universal Value (OUV) (Jokilehto 2008, UNESCO 2008). Among the six OUV selection criteria for cultural heritage, Rome is justified with five of them, including a large variety of heritage values and attributes (Tarrafa Silva et al., 2010, Veldpaus, 2015): Criterion (i) about its artistic/architectural masterpiece, Criterion (ii) about its decisive influence, Criterion (iii) about the archaeological testimony of the civilization, Criterion (iv) about its architectural/monumental typologies, and Criterion (vi) about religious and cultural associations.

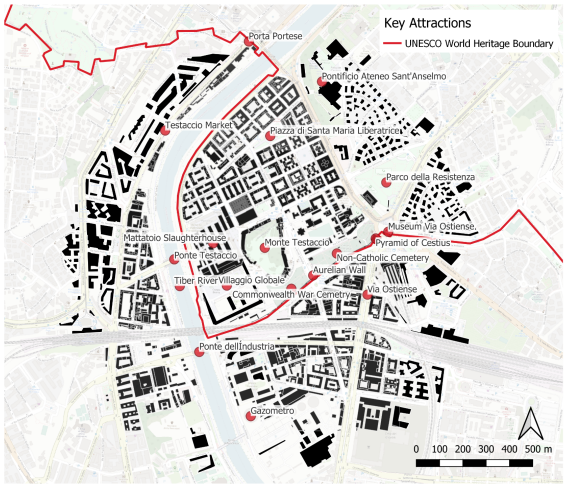

Figure 1. Major tourist attractions labelled in the case study: the Testaccio area in Rome, Italy. The UNESCO World Heritage boundary is marked as a red line, the northeastern side of which is within the World Heritage property.

In Testaccio, specifically, the archaeological excavations and built heritage of great chronological and typological diversity make up the typical urban character of the area (Segarra Lagunes, 2004, De Kleijn et al., 2013, Pica et al., 2016, Burgers et al., 2018, Zheng 2023), including but not limited to Pyramid of Cestius (18 BC), Aurelian Wall (3rd century AD), the Tiber River, Monte Testaccio with winemaker's warehouses (which currently hosts several night-life avenues), the Non-Catholic Cemetery (where, among others, the English romantic poets, John Keats and Percy Bysshe Shelley rest), Commonwealth War Cemetery, and Mattatoio Slaughterhouse (which is currently part of the campus of Roma Tre University). An overview of the major attractions of the area including the boundary of UNESCO World Heritage property is shown in Figure 1 with QGIS (QGIS Development Team, 2023), where the context is drawn based on the OpenStreetMap (Haklay et al., 2008).

2.2 Data Collection and Pre-processing

An image-sharing social media platform (Flickr) is used to collect user-generated information in the Testaccio area. Following the workflow proposed earlier (Bai et al., 2022), the Flickr API is used to collect a maximum of 3000 post IDs with a query of 500m radius centred at 41.87580N 12.47515E. Among the collected 3000 post IDs, only 2040 posts have been indicated as "*candownload*" by the posting owners, which are further used to collect the original posts containing the small-sized (150×150 px) images, the corresponding textual descriptions (might be empty), the geo-locations of the posts, and the registered locations of the users. User information and privacy are fully respected, and the users are pseudonymized. The users whose registered location contains "*rome*" or "*roma*" are considered locals, while the others are considered tourists.

The textual descriptions of the posts, if any, are first split into sentences. The *langdetect* Python library is used to detect the original language of each sentence and only keep the English (513 sentences) and Italian (525 sentences) ones. The *Google Translator* API from the *Deep Translator* Python library is used to translate Italian sentences into English since the models used in this paper are pre-trained with English corpus. After further cleaning the sentences following Bai et al. (2022), 436 English and 503 Italian sentences are obtained to be further analysed.

2.3 Data Analysis with AI Models

Several pre-trained deep learning models in the emerging fields of computer vision and natural language processing (NLP), and pre-defined dimensionality reduction algorithms are applied in different steps of this study for the analyses.

Inspired by Dombrowski (2019), the images are processed with *VGG-16* network (Simonyan and Zisserman, 2015) pre-trained on the ImageNet dataset (Deng et al., 2009) using the *Keras* Python library to obtain the last 4096-dimensional (4096D) hidden layer vector output as their structured data representations. The 4096D visual features are further reduced to 300D vectors with Principal Component Analysis (PCA). The PCA features are then fed into a *t*-distributed Stochastic Neighbour Embedding (*t*-SNE) algorithm using *Scikit-Learn* to make a 2D embedding of the features (Van der Maaten and Hinton, 2008). The images are then visualised by transforming the *t*-SNE coordinates of the data points into a regular 2D grid using the *RasterFairy* Python library. As such, the original unstructured complex social media images are simplified as a 2D representation, where closer images in this 2D space imply shorter distances of them in the high-dimensional feature space, thus more visually similar based on their depicted scenes. For exploratory and demonstrative purposes, 1000 images randomly sampled from the image collection are visualized, which are eventually clustered manually with their main depicted topics. The identified representative visual clusters could be interpreted as proxies for the perceived and expressed urban heritage attributes by social media users.

Moreover, the pre-processed sentences mentioned in Section 2.2 are fed into two state-of-the-art NLP models (i.e., *BERT* and

ULMFiT) pre-trained on a dataset of UNESCO World Heritage Statements of OUV (Bai et al., 2021b). The relevance of each sentence to the ten OUV selection criteria is predicted by the models as probability distributions. Since not all sentences are relevant to cultural heritage, possibly confusing the NLP models trained therefrom, the consistency (the Intersection over Union of top-3 predictions) and confidence (the sum of top-3 logits) of the two models are used to filter the results (Bai et al., 2022). After filtering, 119 sentences (68 English and 51 Italian) from 53 unique posts by 28 users remain. The OUV selection criterion as top-1 prediction by the models for each sentence is considered as its label of perceived and expressed urban heritage values, similar to the visuals. Furthermore, word clouds containing both English and Italian terms are generated using the *WordCloud* Python library with the original untranslated sentences classified as each OUV selection criterion.

The geographical locations of all the collected posts owned by both locals and tourists are visualized using QGIS, with which a kernel density heatmap of posts is also generated using a 200-metre radius. Similarly, those posts containing sentences that are classified as relevant to the OUV selection criteria are also mapped in QGIS as coloured dots.

3. RESULTS

3.1 Distributions of Social Media Data

Even though only posted on Flickr in the digital era, the collected images were originally taken in the range from 1988 to 2021, while the majority (95.7%) of images were taken between 2015 and 2021. A total of 342 unique users contributed to the discussion arena on Flickr in the Testaccio area, among which 63 are identified as locals and 279 as tourists following the procedure mentioned in Section 2.2.

Figure 2 visualizes the geographical distribution of the posts owned by different groups of people. It proves that from a bottom-up perspective, both local Roman people and tourists from all over the world are actively present and eager to share their observations and experiences they have in this area on Flickr. Comparing the heatmap in Figure 2 to the official heritage and major attractions in Figure 1, it is evident that places such as the Pyramid, the non-Catholic cemetery, the Mattatoio ex-Slaughterhouse, and the ex-warehouse (current bars) around Monte Testaccio all attract a significant amount of interest, showcasing a consistent representation of the urban images revealed on social media platform with the official heritage and attractions. However, two other "unexpected" significant hotspots appear in Figure 2 that are not initially marked in Figure 1: one in the northwestern corner of the study area near the Testaccio market, and the other in the south near the railway. Further investigations of the data show that the former is due to a large number of images posted by the same user at the same spot depicting a large-scale get-together and that the latter is due to a broad variety of street art and graffiti near Porto Fluviale. Both patterns will also be present in the visualization of post images and topics in Figure 3.

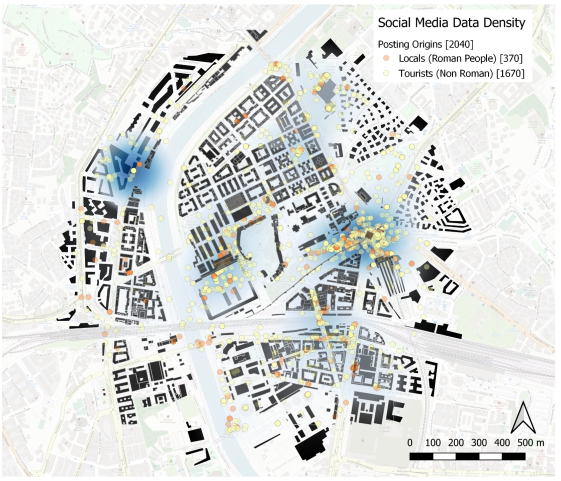

Figure 2. The distribution of social media images collected in the area of Testaccio. The posts from locals and tourists are overlaid with a heatmap of all posts.

3.2 Depicted Elements as Heritage Attributes

Figure 3 demonstrates the identified topics of 1000 sampled images in the area. Images posted by locals are kept polychrome while those posted by tourists are turned into monochrome.

Figure 3. Social media image clusters based on the sampled image content using *t*-SNE algorithm with their respective proportions.

The plot with transformed *t*-SNE embeddings significantly eases the manual clustering process, as the images with similar visual content are generally placed close to each other, which is evident in the cases of Pyramid in the top part of Figure 3, the Gazometro in the top-right corner, human portraits in the bottom, and urban scenes that span in the middle. This visualization provides a zoomed-in view and interpretation of the geographical distribution in Figure 2: not only where people frequently post (i.e., the context), but also what people frequently depict in those images (i.e., the content). It shows that most places mentioned in Section 2 can be observed in the online gallery of Flickr. The visualized images in Figure 3 are a random sample of the collected dataset in this study, which itself is a [semi-]random sample of all the Flickr posts and a further abstraction and simplification of the real world. Still, the proportion of the identified topic clusters here can be assumed to show certain representativeness (even though this could also be probably skewed, which needs to be further backed up and corrected with other sources of data and types of research, as will be discussed in Section 4.2). According to this visualization, the most dominant elements of the area are the Pyramid (8%), the Cemetery (12%), Street Art and Graffiti (13%), and the nightlife along the Testaccio market and around Monte Testaccio (25%). Note that many of the "NightLife and Gastronomy" images are posted by a single user at a single spot, as pointed out in Section 3.1, suggesting that this "largest cluster" might not be a reliable estimation. Still, together with other similar images, this cluster shows a strong sign that people actively live in the area with various activities. Further discussions on the alternative treatment of data skewness can be found in Section 4.2. Interestingly, pictures of Street Art and Graffiti with remarkably high quality and vividness take up a significant proportion of the results, suggesting that this element could be a typical characteristic of this area representing its modern urban images, together with other conventional heritage properties and attractions, which could be further emphasized in future planning, as will be argued in Section 4.1.

Except for the presence, the absence of certain visual elements in the search can also provide useful messages. For example, an issue of interest revealed with the visualisation is that the visual representations of Monte Testaccio, the Aurelian Wall, and the Tiber River are not significant on Flickr, suggesting that these formal heritage attributes are not given enough attention by the sampled Flickr images. This may be due to various reasons, including but not limited to the accessibility and visibility problems of certain places, the awareness of the heritage values and attributes, the specific interests of the group of Flickr users (which are mainly photographers), and the under-representativeness during the random sampling procedure. The first assumption has been further confirmed with site analyses in Section 4.1, where a few planning proposals for improvement are also introduced.

3.3 Described Topics in Support of OUV

Figure 4 shows the distribution of the sentences from Flickr posts classified by the NLP models (with relatively high confidence and consistency) as strongly related to various OUV selection criteria, thus showing heritage values and attributes conveyed to the post owners. All five OUV selection criteria of the Roman UNESCO World Heritage property could be observed as perceived by [some of] the online community in the area, while Criterion (vi) about association, Criterion (iii) about testimony, and Criterion (i) about masterpiece are the most representative characteristics.

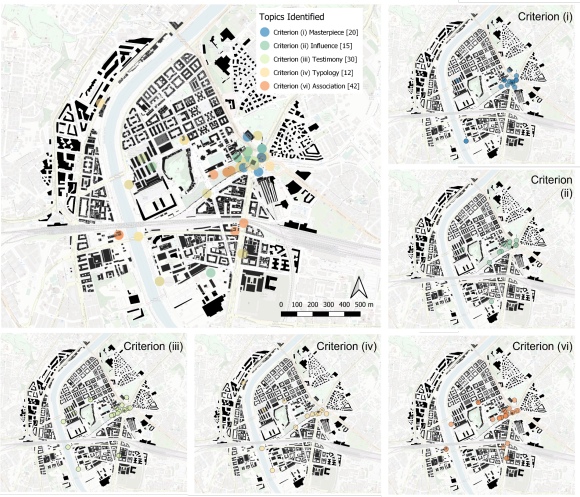

Figure 4. The distribution of the sentences classified to be relevant to the OUV selection criteria.

The OUV-related posts are mostly concentrated in the area of the Pyramid, the Non-Catholic Cemetery, and along the Via Ostience, the Tiber River, and the Aurelian Wall. This demonstrates a similar pattern visible in Figure 3, while the "surprising" elements of the party places are not captured. This is partly because the NLP models used in this study are originally trained with the official Statements of OUV by UNESCO, thus effectively mainly a machine replica of the expert-based view. Specifically, descriptions about Criterion (i) - Masterpiece can be found mainly around the Pyramid and at the ex-Gazometro, that about Criterion (ii) - Influence mainly along the Aurelian Wall including the San Paolo gate, that about Criterion (iii) - Testimony mainly in the Non-Catholic Cemetery, that about Criterion (iv) - Typology along the buildings, chapels, and also the "*long, linear market*" in the area, and finally, that about Criterion (vi) - Association in all places mentioned above describing human activities. It is worth noting that one single post (usually long ones posted by enthusiastic Flickr users) can possibly contain multiple sentences that are classified as relevant to various OUV selection criteria. This is evident in a rather long post introducing the history of the non-Catholic cemetery and John Keats, classified as related to Criteria (ii), (iii), and (vi). Interestingly, another long post emerges talking about the "*street art works with high impact*" in this area, dedicated to the poet, writer, and director Pier Paolo Pasolini, "*one of the greatest Italian artists and intellectuals of the twentieth century*". This post contains sentences classified as related to Criteria (i), (ii), (iii), and (vi), giving another layer of meaning to these popular elements.

Moreover, in the word clouds generated for the three most common OUV selection criteria (i), (iii), and (vi) in the area as a collection of both English and Italian words, as shown in Figure 5, interesting patterns can be discovered. For the sentences classified as related to Criterion (i) - Masterpiece, words concerning both tangible attributes including the names of attractions (such as the *pyramid* and the *cemetery*) and building elements (such as *sculpture*, *brick*, *marble*, *concrete*, *etc*.), and intangible attributes (such as *designed*, *built*, *story*, *etc*.) are frequently present. In sentences classified as concerning with Criterion (iii) - Testimony, most

words are attributed to the elements in the Non-Catholic cemetery, yet pointers to the historic/age values such as the word *ancient* and several historic dates (e.g., *Marzo*, *CCCXXX*) appear. Yet in sentences classified as relevant to Criterion (vi), while the main topics are still about the Pyramid and the Cemetery, additional focuses are given to names of well-known historic figures (such as *Keats*, *Shelley*), religious terms (such as *pellegrinaggio*, *catholic*), and cultural origins (such as *Inglesi*, *Giapponese*, *Bulgaro*, *Lituano*, etc.). Note the following points: 1) only single words rather than N-grams (continuous phrases) have been included in the word cloud generation process, which could be improved in the future; 2) some words appear repeatedly in all generated word clouds (such as *roma*, *pyramid*, *cemetery*), but they are not necessarily unimportant “stop words” that could be simply filtered out, which rather show the complex and multi-layered nature of cultural heritage; 3) not all words involved in the word clouds are meaningful and/or informative, which should be further polished by heritage and planning experts, as sentence classification task is a collaboration of the pre-trained NLP models and the actual collected data in the study area. Therefore, the behaviour shown in the word clouds strongly relies on the quality of both components.

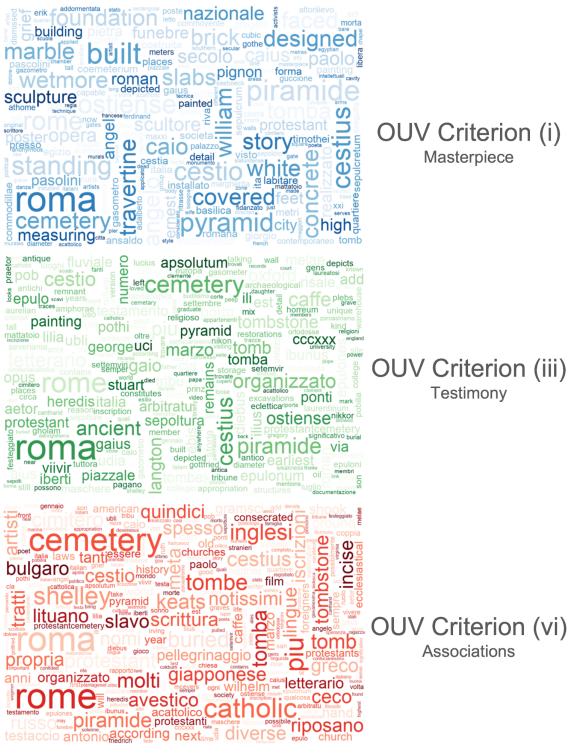

Figure 5. Word clouds generated with posts classified as relevant to one of the three most significant OUV selection criteria in the study area. The larger a word, the more frequently it appeared in the corpus.

4. DISCUSSION

4.1 Implication for Heritage and Urban Planning

Under the principles of Historic Urban Landscape (HUL), the collected data can contribute to integrating the public perspective into urban heritage planning and management. From a heritage planning perspective, several hypotheses can be drawn for the valorisation of the local heritage in the Testaccio neighbourhood, which could also become indicators for improving accessibility, liveability, and connectivity of the area. Figure 3 shows the dominance of some elements (such as the Pyramid, the Cemetery, the graffiti and street art, and the nightlife), and a lack of attention to some other elements (such as Monte Testaccio, the Aurelian Wall, and the Tiber River) on Flickr. This could suggest that some of the formal heritage sites and attributes (Monte Testaccio, the Aurelian Wall, as well as the Tiber River) are not well-perceived by residents and tourists. Further on-site observations and interviews suggest the reasons could be the lack of physical accessibility, visibility, and/or a sense of safety. At the same time, street art has almost the same level of social media publicity from both residents and tourists as the dominant heritage landmarks such as Pyramid and Cemetery, which could imply that informal and spontaneous cultural elements are equally valuable for people as the authorised heritage.

The analyses of social media data and AI models were integrated with other sorts of information from primary sources, such as interviews with local stakeholders, direct observation, and field surveys, as well as with the study of state-of-the-art policy frameworks, planning and designing documentation, and existing literature of previous research. The integration of the information from these various sources largely agreed with the outcomes of social media analysis. Namely, three main aspects were confirmed: the lack of connectivity throughout the neighborhood due to physical barriers such as walls and fences all around the area; the lack of safety, security, and connectivity along the banks of the Tiber River, where only two access points to the river remain in the neighborhood; and the lack of access to some specific heritage sites. For example, Monte Testaccio is nowadays only a landmark of “present absence”, as most of the time access to it is closed to both inhabitants and tourists. In addition to formal heritage attributes, the integrated analysis also confirmed a thriving street art culture in the area, which is being used to requalify derelict and/or abandoned buildings and physical barriers in the surroundings. The interviews also pointed to the important role of Mattatoio, the former slaughterhouse that appears to be actively used as a pedestrian passage between the riverfront and the Testaccio neighborhood, and also as a hub for local cultural and social actors such as the *Città dell’Altra Economia*, the *Associazione Culturale Roma Open Lab,* and *L’altro Abitare*, working on the circular economy, traditional art and craft, and refugees hospitality. This feature is further amplified by the presence of the Roma Tre University in Mattatoio, which pulls in the young generations to the area. Furthermore, the interviews also revealed that the neighborhood has an active network of local community associations, such as *Testaccio in Testa*, *Rione Testaccio*, and the *Condominio* informal library, most of which have initiated bottom-up participatory activities over the years, developing planning suggestions to improve the neighbourhood.

The integrated analyses of issues and potentials assisted in developing tailored planning suggestions to address the identified issues. It was made clear that the issues can be dealt with through participatory processes and local small-scale improvements, such as by the urban acupuncture or tactical urbanism approaches (Lydon et al., 2015; Silva, 2016) without top-down capital-driven investments. The existing active network of local associations carries the potential to make such participatory actions effective and sustainable in the long term. Some of such soft interventions

could, for example, dwell on the existing popular street art culture in the area. A "street museum" could bring the community together in proposing and implementing locally inspired art projects, turning the existing physical barriers: walls, fences, and crossroads into communal spaces. Along with the art projects, local community organisations could lead small-scale improvements in the public spaces, taking ownership of the degraded public spaces and nurturing long-term community attachment and care for these areas. The local community could play a particularly important role in the careful revitalisation of riverbanks through cleaning-up campaigns and the installment of recreational facilities, such as those for outdoor sports, bird watching, ecological education, etc. More active and regular use of the area would contribute to a safer space and encourage still more diverse sustainable uses of the territory. Based on the local consultations, other more conventional urban interventions could also be developed, such as a pedestrianised piazza in front of Matattoio at the Testaccio bridge and Piazza Testaccio, promoting pedestrian use and greening public spaces within the Mattatoio premises. The careful redesign and/or improved management of traffic flows and pedestrian crossings could provide a public space with better quality, accessibility, and connectivity through the area towards the riverfront. Although these are only proposals, they show how social media data collection and analyses can be employed in participatory planning and decision-making in the complex urban heritage realm. More importantly, it also demonstrates that participation should not be limited to data collection only (Pereira Roders, 2019). A participatory process should involve all various stages of sustainable improvements, from early consultations to design, decision-making, and implementation, stimulating the local community to set up solutions for their living environment and providing enabling conditions for implementing them. In the case of Testaccio, such an approach could benefit the community organisations that are already active in the area, and transform them into a social network capable of taking ownership of local management and improvement actions.

In summary, the social media analyses supported with AI models can potentially become an important tool for future cultural heritage planning and urban design projects under the recommendation of HUL, since they gather a broad range of information from a diverse group of actors. Issues and potentials that emerge in the analyses could inform and inspire the participatory planning processes. Social media could also enable observations of how the current ongoing planning actions could gradually alter first the physical spaces and then the "digital twin" of the study area by collecting freshly-added posts by residents and tourists along the timeline and merging them into the same dataset. This new collection of data will be processed and analysed with newer generations of more powerful AI algorithms in the future. Then a new round of research could be conducted, coupled with integrated analyses of mixed methods, possibly again pointing to new planning directions. This would suggest an abductive and iterative system with a data-driven feedback mechanism for decision-making, integrating a diversity of data sources in an inclusive and participatory planning process (Dubois et al., 2002, Pereira Roders, 2019).

4.2 Limitation and Future Steps

The methodology of mapping urban heritage images using social media data and artificial intelligence proposed in this paper mainly serves exploratory analyses and demonstrative purposes. Even though strongly aided by AI models and algorithms, several components of the workflow, e.g., the manual clustering of images based on their visual similarity, still rely significantly on experts in the loop. This could be easily updated with other [semi-]automatic clustering methods such as k-Means and DBSCAN (Wu 2012, Schubert et al., 2017). The number of users involved, posts analysed, and images visualised in this study ended up rather small, making the analyses also manageable by hand. However, as the scale of the case study further grows in application scenarios elsewhere other than Testaccio, the necessity and benefit of an AI-aided procedure would become more trivial, especially considering the systematic, efficient, scalable, and generalisable performance of various pre-trained AI models. Even though only a very limited number of posts are found to be clearly related to one or more of the OUV selection criteria, thus significantly describing topics related to heritage values and attributes, the results do not necessarily disqualify the cultural significance of the Testaccio area. This is on one hand, due to the data source chosen in this study, since Flickr is mainly an image-sharing platform, thus people seldom write extensive essays; on the other hand, due to the rather strict filtering rules applied in this study based on the confidence and consistency of the two pre-trained models. Multi-modal machine learning techniques fusing the visual and textual information together with attention mechanisms during the training and inference could significantly improve the precision and recall of valuable information concerning cultural significance (Baltrušaitis et al., 2018, Bai et al., 2022). Only a small fraction of AI models are tested as a humble exploration in this study. Much more possibilities and potentials still exist for applying AI in social media mining and in heritage management. However, even with a fully automatic workflow at a later stage, inspections and interpretations of the results by humans with their expert knowledge are always needed.

The results in Figures 2 and 3 showed that if one single user posted too many redundant pictures at the same place, the distribution could be significantly skewed. This effect could be easily fixed by filtering out the redundant posts, such as in Ginzarly et al. (2019). However, multiple posts at the same place by the same user do not necessarily mean a redundancy in information, which might also indicate strong interests by a specific group of people. Simply removing "redundant" posts might also obscure the abstraction of reality. Future studies are suggested to balance between the two extremes, to find the tailored configuration for their application scenarios, and to make suitable interpretations accordingly.

Nevertheless, the use of AI and social media data is never the "eternal solution" for mapping urban heritage images, which could create more new challenges and problems than it manages to solve. Furthermore, an important limitation of the use of this method in support of participatory planning is the unequal representation of users and non-users. As broader research demonstrates, despite the proliferation of digital technologies, a significant number of the population may still be disadvantaged in using digital platforms and tools, due to a lack of access to the internet, equipment, and difficulty with digital skills (Craglia et al., 2021). Some of these inequalities are related to age ranges and socio-economic backgrounds or in other cases to spatial divides, for example, lack of internet coverage in remote rural or natural areas. Thus the outcomes of social media surveys may be considered unavoidably biased towards the users of digital platforms, implying a generational, socio-economic, and/or spatiotemporal gap in

its representation. Other factors for consideration may be the changing engagement modes and attitudes over time, competition between different platforms, or the risks associated with personal data protection. These factors call for careful consideration at the early stage of planning and emphasize the need for integrated research and mixed analysis methods combining qualitative and quantitative knowledge. Social media may be helpful for setting the stage for planning through an initial data capture, but its limitations should also be balanced with other methods of data collection and analyses as well as cross-sectoral integration of different data sources, such as interviews, surveys, observations, and participatory workshops (Zheng, 2023).

5. CONCLUSION

This paper presented an exploratory case study using artificial intelligence to map the perceived urban heritage images with user-generated social media data of the Testaccio area in Rome. The proposed methodology provided an alternative perspective for viewing the urban heritage as a collection of generated content, which is to be eventually integrated into a series of heritage-related investigations of mixed methods. Despite its limitations, this semi-automatic workflow can serve as a documentation tool contributing to the collective view and knowledge of urban cultural heritage, summarizing and highlighting the opinions and attention of the public. Following the recommendation of Historic Urban Landscape, this collective knowledge can lead to potentially more inclusive heritage management that can fertilize the process of local development planning.

ACKNOWLEDGEMENTS

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 Sklodowska-Curie grant agreement No 813883.

REFERENCES

Assmann, J., & Czaplicka, J., 1995. Collective memory and cultural identity. *New German critique*, (65), 125-133.

Bai, N., Nourian, P., & Pereira Roders, A., 2021a. Global citizens and world heritage: Social inclusion of online communities in heritage planning. The *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 46, 23-30.

Bai, N., Luo, R., Nourian, P., Pereira Roders, A., 2021b. WHOSe Heritage: Classification of UNESCO World Heritage statements of ”Outstanding Universal Value” with soft labels. *Findings of EMNLP 2021*, Association for Computational Linguistics, Punta Cana, Dominican Republic, 366–384.

Bai, N., Nourian, P., Luo, R., Pereira Roders, A., 2022. Heri-Graphs: A Dataset Creation Framework for Multi-Modal Machine Learning on Graphs of Heritage Values and Attributes with Social Media. *ISPRS International Journal of Geo-Information*, 11(9).

Bai, N., Nourian, P., Pereira Roders, A., Bunschoten, R., Huang, W., & Wang, L., 2023. Investigating rural public spaces with cultural significance using morphological, cognitive and behavioural data. *Environment and Planning B: Urban Analytics and City Science*, 50(1), 94-116.

Baltrušaitis, T., Ahuja, C., & Morency, L. P., 2018. Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2), 423-443.

Batty, M., 2023. A new kind of search. *Environment and Planning B: Urban Analytics and City Science*, 50(3), 575-578.

Boy, J. D., & Uitermark, J., 2017. Reassembling the city through Instagram. *Transactions of the Institute of British Geographers*, 42(4), 612-624.

Burgers, G. J., Contino, A., D’Alessandro, L., De Leonardis, V., Della Ricca, S., Kok-Merli, R. A., & Sebastiani, R., 2018. The afterlife of the Porticus Aemilia. *Journal of Fasti Online*, (400), 1-19.

Craglia, M., Micheli, M., Hradec, J., Calzada, I., Luitjens, S., Ponti, M., ... & Boter, J., 2021. *Digitranscope: The governance of digitally-transformed society.*

Cho, N., Kang, Y., Yoon, J., Park, S., and Kim, J., 2022. Classifying tourists’ photos and exploring tourism destination image using a deep learning model. *Journal of Quality Assurance in Hospitality & Tourism*, 1–29.

De Kleijn, M., van Aart, C. J., Van Manen, N., Burgers, G. J., Scholten, H. J., 2013. Testaccio, a digital cultural biography app. *UMAP Workshops*.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In 2009 *IEEE conference on computer vision and pattern recognition* (pp. 248-255).

Dubois, A., & Gadde, L. E., 2002. Systematic combining: an abductive approach to case research. *Journal of business research*, 55(7), 553-560.

Ducci, M., Janssen, R., Burgers, G. J., & Rotondo, F., 2023. Mapping Local Perceptions for the Planning of Cultural Landscapes. *International Journal of E-Planning Research (IJEPR)*, 12(1), 1-27.

Estellés-Arolas, E., & González-Ladrón-de-Guevara, F., 2012. Towards an integrated crowdsourcing definition. *Journal of Information science*, 38(2), 189-200.

Fostikov, A., 2023. First impressions on using AI powered chatbots, tools and search engines: ChatGPT, Perplexity and other–possibilities and usage problems.

Galesic, M., Bruine de Bruin, W., Dalege, J., Feld, S. L., Kreuter, F., Olsson, H., … & van Der Does, T., 2021. Human social sensing is an untapped resource for computational social science. *Nature*, 595(7866), 214-222.

Ginzarly, M., Roders, A. P., & Teller, J., 2019. Mapping historic urban landscape values through social media. *Journal of Cultural Heritage*, 36, 1-11.

Haklay, M., & Weber, P.,2008. OpenStreetMap: User-generated street maps. *IEEE Pervasive computing*, 7(4), 12-18.

Janssen, J., Luiten, E., Renes, H., and Stegmeijer, E., 2017. Heritage as sector, factor and vector: conceptualizing the shifting relationship between heritage management and spatial planning. *European Planning Studies*, 25(9): 1654–1672.

Jokilehto, J., 2008. *What is OUV? Defining the Outstanding Universal Value of Cultural World Heritage Properties*. Technical report, ICOMOS.

Kang, Y., Cho, N., Yoon, J., Park, S., Kim, J., 2021. Transfer learning of a deep learning model for exploring tourists' urban image using geotagged photos. *ISPRS International Journal of Geo-Information*, 10(3), 137.

Kumar, P., 2020. Twitter, disasters and cultural heritage: A case study of the 2015 Nepal earthquake. *Journal of Contingencies and Crisis Management*, 28(4), 453-465.

Lai, J., Cheng, T., & Lansley, G., 2017. Improved targeted outdoor advertising based on geotagged social media data. *Annals of GIS*, 23(4), 237-250.

LeCun, Y., Bengio, Y., & Hinton, G., 2015. Deep learning. *Nature*, 521(7553), 436-444.

Lefebvre, H, 1991. The production of space. In *The people, place, and space reader* (pp. 323-327). Routledge.

Liu, P., & De Sabbata, S., 2021. A graph-based semi-supervised approach to classification learning in digital geographies. *Computers, Environment and Urban Systems*, 86, 101583.

Lydon, M., Garcia, A., Lydon, M., & Garcia, A., 2015. *A tactical urbanism how-to* (pp. 171-208). Island Press/ Center for Resource Economics.

Lynch, K., 1964. *The image of the city*. MIT press.

Monteiro, V., Henriques, R., Painho, M., & Vaz, E., 2014. Sensing World Heritage: An Exploratory Study of Twitter as a Tool for Assessing Reputation. In *Computational Science and Its Applications–ICCSA 2014: 14th International Conference*, Guimarães, Portugal, June 30–July 3, 2014, Proceedings, Part II 14 (pp. 404-419).

Pan, S. J., & Yang, Q., 2010. A survey on transfer learning. IEEE *Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Pereira Roders, A., 2019. The Historic Urban Landscape approach in action: Eight years later. In *Reshaping urban conservation: The historic urban landscape approach in action* (pp. 21-54). Singapore: Springer Singapore.

Pica, A., Vergari, F., Fredi, P., & Del Monte, M., 2016. The Aeterna Urbs geomorphological heritage (Rome, Italy). *Geoheritage*, 8, 31-42.

QGIS Development Team, 2023. QGIS Geographic Information System, Open Source Geospatial Foundation, http://qgis.osgeo.org.

Schich, M., Song, C., Ahn, Y. Y., Mirsky, A., Martino, M., Barabási, A. L., & Helbing, D., 2014. A network framework of cultural history. *Science*, 345(6196), 558-562.

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X., 2017. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, 42(3), 1-21.

Segarra Lagunes, M. M., 2004. *Il Tevere e Roma: Storia di una simbiosi*. 1-432.

Silva, P., 2016. Tactical urbanism: Towards an evolutionary cities' approach?. *Environment and Planning B: Planning and design*, 43(6), 1040-1051.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. Y. Bengio, Y. LeCun (eds), *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015.*

Tarrafa Silva, A. and Pereira Roders, A., 2010. The cultural significance of World Heritage cities : Portugal as case study. In *Heritage and Sustainable Development*, pages 255–263, Évora, Portugal.

Thorp, H. H., 2023. ChatGPT is fun, but not an author. *Science*, 379(6630), 313-313.

UNESCO, 2008. Operational guidelines for the implementation of the world heritage convention. Technical Report July, UNESCO World Heritage Centre.

UNESCO, 2011. Recommendation on the Historic Urban Landscape. Technical report, UNESCO World Heritage Centre.

Van der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Veldpaus, L., 2015. *Historic urban landscapes: framing the integration of urban and heritage planning in multilevel governance.* PhD thesis, Technische Universiteit Eindhoven.

Watkins, J., 2007. Social media, participatory design and cultural engagement. In *Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces* (pp. 161-166).

Wu, J., 2012. Cluster analysis and K-means clustering: an introduction. *Advances in K-Means clustering: A data mining thinking*, 1-16.

Zhang, F., Zhou, B., Ratti, C., & Liu, Y., 2019. Discovering place-informative scenes and objects using social media photos. *Royal Society open science*, 6(3), 181375.

Zheng, N., 2023. *Coming to grips with diverse voices in participatory heritage initiatives*. PhD Thesis, Vrije Universiteit Amsterdam.