



Exploring Public Perception of Urban Heritage in Macao via User-Generated Content

GEO 511 Master's Thesis

Author: Yilin Wei, 23-742-570

Supervised by: Dr. Inhye Kong

Faculty representative: Prof. Dr. Ross Purves

29.01.2026



**Universität
Zürich**^{UZH}

GEO 511 Master Thesis
presented to the Faculty of Science
of the University of Zurich
for the degree of
Master of Science in Geography

Exploring Public Perception of Urban Heritage in Macao via User-Generated Content

Author: YILIN WEI
Student ID: 23-742-570

Supervisor: Dr. Inhye Kong
Faculty Member: Prof. Dr. Ross Purves

Department of Geography

Submission Date: 26.01.2026

Abstract

This thesis examines how the Historic Centre of Macao is described, evaluated, and visually represented in user-generated content on major travel and map platforms. The aim is to understand how visitors express their experiences of the site in everyday language and images, and how these public narratives relate to the Outstanding Universal Value (OUV) that justifies its inscription under criteria (ii), (iii), (iv), and (vi). The study adopts a place-based framework at the level of the 27 officially designated heritage components. Multilingual reviews from Google Maps, TripAdvisor, and Ctrip (English, Simplified Chinese, Traditional Chinese) are analysed using sentiment classification, semantic keyterm extraction, and a two-stage OUV alignment pipeline. On the visual side, user-uploaded images are embedded with a pretrained vision model and grouped into global visual themes, which are then linked to textual and OUV profiles.

The results show that reviews of the Historic Centre are strongly positive overall and focus on churches, squares, facades, gardens, walking, and taking photos, portraying the site as a lived urban landscape. OUV alignment reveals a structured but uneven correspondence with official criteria: Criterion (iv) (material urban fabric and architectural ensembles) dominates, Criterion (iii) (cultural fusion) is secondary, while Criteria (ii) and (vi) are weakly expressed. At the same time, user-generated content highlights concerns outside the OUV framework, such as crowding and commercialisation. Components visually dominated by landmark facades and streetscapes tend to have higher shares of Criterion (iv) claims and more positive sentiment, indicating that textual and visual signals are partially overlapping but not redundant. The thesis contributes a multi-modal method for linking public narratives to OUV frameworks and offers insights that can support monitoring and communication for the Historic Centre of Macao and other urban World Heritage sites.

Acknowledgment

I would like to express my sincere gratitude to everyone who supported me throughout the completion of this thesis.

First and foremost, I am deeply grateful to my supervisors, Prof. Dr. Ross Purves and Dr. Inhye Kong, for their guidance, patience, and encouragement. Their feedback helped me sharpen the research framing, improve the methodological clarity, and—most importantly—keep moving forward when the project became complex. I truly appreciate the trust Dr. Inhye Kong placed in my ideas and the steady support she offered as the thesis evolved from a broad interest in digital heritage discourse into a structured, multi-modal analysis.

I would also like to thank my classmates in the GIScience in the University of Zurich and the GIScience programme for providing an inspiring space to learn and work. Discussions, informal advice, and the shared culture of careful research made this process feel less solitary and more collaborative.

This research relies on publicly available user-generated content from major travel and map platforms(e.g., Google Maps, TripAdvisor, and Ctrip). I acknowledge the platforms and the communities of contributors whose reviews and images made this study possible. While the thesis analyzes these materials in aggregate for research purposes, I remain aware that each data point reflects a personal experience and a particular moment in the site’s living digital history.

I also acknowledge the external tools and resources that supported this thesis. I used OpenAI’s ChatGPT for brainstorming and debugging support, while all interpretations and remaining errors are my own. Parts of the workflow were developed and run in Google Colab. Google Translate was used to assist the translation and standardisation of key terms during exploratory analysis. I also drew on open resources published by the Macao SAR Government for contextual grounding. All external sources and tools used are explicitly cited in the thesis.

Finally, I am grateful to my friends and family for their understanding and support. Their encouragement, in many small but meaningful ways, helped me sustain the focus needed to bring this thesis to completion.

Contents

Abstract	ii
Acknowledgment	iii
List of Acronyms	viii
1. Introduction	1
1.1. Background	1
1.2. Motivation	3
1.3. Research Questions	5
1.4. Thesis Structure	6
2. Related Work	7
2.1. Public Perception and Urban Heritage	7
2.1.1. UNESCO World Heritage and OUV	7
2.1.2. Heritage Interpretation and Public Perception	8
2.1.3. Expert Narratives vs. Public Narratives	9
2.2. Tourism, Social Media Platforms and UGC	11
2.2.1. UGC in Tourism and Destination Image	11
2.2.2. Platforms as Different Tourist Cultures	13
2.2.3. Heritage-Specific Uses of Social Media	14
2.3. Language and Multilingual Public Perception	15
2.3.1. Language as a Lens of Culture and Heritage Meanings	15
2.3.2. Multilingual Social Media in Tourism	15
2.3.3. Translation and Cross-Language Comparison	16
2.4. NLP and Computer Vision for Tourism/Heritage	17
2.4.1. Text Analytics in Tourism and Heritage	17
2.4.2. Sentence Embeddings and Claim Alignment	19
2.4.3. Vision Methods in Tourism and Heritage	19
2.4.4. Strengths and Pitfalls of NLP and Vision Methods	20
3. Materials	22
3.1. Overview	22

3.2. Case Study: Macao World Heritage	22
3.3. Textual User-Generated Content Collection	24
3.3.1. Google Maps	24
3.3.2. TripAdvisor	25
3.3.3. Ctrip	26
3.4. Visual User-Generated Content Collection	26
3.5. UNESCO Documentation and OUV Knowledge Base	27
3.6. Data Cleaning and Preprocessing	28
3.6.1. Textual Data	29
3.6.2. Metadata Alignment	30
3.6.3. Image Filtering	30
3.6.4. Data Summary	31
3.7. Legal Considerations and Limitations	32
3.7.1. Ethical and Legal Considerations	32
3.7.2. Limitations of the Materials	32
4. Methodology	33
4.1. Research Design and Analytical Overview	33
4.2. Place-Based Data Organization and POI Definition	34
4.3. Textual Analysis I: Sentiment and Keyterms	35
4.3.1. Sentiment Analysis	35
4.3.2. Keyterm Extraction	36
4.3.3. Cross-Platform and Cross-Language Experiential Profiles	37
4.4. Textual Analysis II: OUV Alignment Pipeline	38
4.4.1. Stage 1: Generic OUV Criterion Classification	38
4.4.2. Stage 2: Macao-Specific OUV Claim Alignment	40
4.4.3. Final Two-Stage Filtering	42
4.4.4. Aggregation and Visualisation	43
4.5. Visual Analysis: Image Embeddings and Clustering	44
4.5.1. Deep Visual Feature Extraction	44
4.5.2. Global Visual Clustering with K-means	44
4.5.3. Qualitative Interpretation and Naming of Visual Themes	46
4.5.4. Visualisation of the Global Feature Space	46
4.5.5. Visual Theme Distributions per Heritage Component	47
4.6. Anchor Case Design: Ruins of St Paul’s	47
4.7. Cross-modal synthesis strategy	50
5. Results	52
5.1. Overview of Outcomes	52

5.2. Anchor Case: Ruins of St Paul's	53
5.2.1. Data Coverage	53
5.2.2. Sentiment and Keyterms	53
5.2.3. OUV-Aligned Sentences	59
5.2.4. Visual Themes	61
5.3. System-wide Textual Patterns	63
5.3.1. Sentiment Across Components, Platforms, and Languages . .	63
5.3.2. Keyterm Profiles and Comparisons	66
5.3.3. OUV Alignment: Global Patterns	73
5.4. System-wide Visual Patterns	77
5.4.1. System-wide Visual Theme Model	77
5.4.2. Visual Theme Definitions and Exemplars	78
5.4.3. Embedding-space Structure of Themes	80
5.4.4. Language-sensitive Preferences	81
5.4.5. Component Visual Theme Profiles	82
5.5. Cross-modal Synthesis	84
5.5.1. Component-level alignment across text and images	84
5.5.2. Summary of Cross-Modal Findings	87
6. Discussion	89
6.1. Revisiting the Research Questions	90
6.2. Public Perception and OUV Alignment	93
6.3. Language and Platform Differences	95
6.4. Cross-modal interpretation: convergence and divergence	97
6.5. Methodological Reflections and Limitations	99
6.5.1. Data sources, Sampling, and Representativeness	100
6.5.2. Textual Pipeline: Sentiment and Experiential Signals	101
6.5.3. OUV Alignment: Modelling Heritage Narratives	102
6.5.4. Visual Themes and Cross-modal Alignment	103
6.5.5. Influences for Interpretation	104
6.6. Implications for World Heritage Management and Future Work . . .	104
7. Conclusion	107
7.1. Summary of Aims and Approach	107
7.2. Synthesis of Main Findings	108
7.3. Contributions, Limitations, and Outlook	109
References	111
A. Additional Tables	124

Personal Declaration

130

List of Acronyms

EN	English
MT	Machine Translation
NLP	Natural Language Processing
OUV	Outstanding Universal Value
POI	Point of Interest
UNESCO	United Nations Educational, Scientific and Cultural Organization
UGC	User-Generated Content
UTF-8	Unicode Transformation Format (8-bit)
XML	eXtensible Markup Language
zh-CN	Simplified Chinese
zh-TW	Traditional Chinese (mainly used in Hong Kong, Macao, Taiwan)

1. Introduction

1.1. Background

Understanding how people perceive and value places is a long-standing concern in geography, tourism studies, and urban research. Places are not only physical settings. They are also experienced, interpreted, and remembered through everyday practices, stories, and representations. In tourism contexts, these interpretive processes become especially visible, because visitors encounter places through expectations, on-site experiences, and post-visit narration. Classic work on sense of place emphasises that place meaning emerges from subjective experience and imagination rather than from formal descriptions alone (Tuan, 1977). In a similar spirit, studies of urban experience highlight that people form mental images of cities based on what they notice, value, and recall (Lynch, 1960). These perspectives motivate empirical attention to public narratives as a way to understand how places are framed and evaluated in practice.

In recent decades, digital media have expanded the channels through which place meanings are produced and circulated. Online platforms have become a central arena for how visitors and residents describe, recommend, and visually portray places. This matters for tourism destinations, where digital information increasingly shapes both decision-making and post-visit interpretation (Xiang and Gretzel, 2010). Digital platforms do not simply store opinions. They also mediate how people talk about place by structuring visibility, prompting specific forms of expression, and embedding reviews and images in platform-specific interface designs and ranking systems. As a result, digital representations of place can be analysed as a distinct layer of contemporary placemaking (Halegoua, 2020).

User-generated content (UGC) has therefore become an important empirical resource for studying public perception at scale. Reviews, ratings, and user-uploaded images offer large volumes of fine-grained, experience-near material that is difficult to obtain through traditional surveys or interviews. This has encouraged a growing body of research using computational approaches to analyse destination image,

visitor experiences, and perceptions expressed in online discourse (Marine-Roig and Anton Clavé, 2016; Bai et al., 2022b; Chen et al., 2024). At the same time, UGC is not a neutral mirror of the population. Participation is uneven, and contributors are typically a self-selected subset of digitally active users. Platform design further shapes what is shared and how it is evaluated. For this reason, cross-platform comparison has become increasingly important in tourism research, because different platforms attract different user groups and promote different norms of review writing and visual posting (Bucher et al., 2018). Comparing platforms can reveal whether observed perception patterns are robust signals or products of a single platform ecosystem.

A second, closely related challenge concerns language. Tourism is inherently multilingual, and places are discussed and discovered in different linguistic registers. Language matters not only as a translation issue but also as a driver of uneven digital coverage and different conceptual framings. Research in geographic information science shows that the availability and structure of web-based geographic information vary across languages and data sources, affecting which places are visible online and how they are described (Venkateswaran et al., 2014). This is particularly relevant for tourism and heritage settings, where international visitors and regional publics engage with the same sites through different platforms and languages. In addition, linguistic differences can shape which place qualities are foregrounded in descriptions. Studies on how people conceptualise landscape terms across languages suggest that meaning structures do not always map directly between linguistic communities, even when people refer to similar physical environments (Purves et al., 2023a). These observations support treating multilingual UGC as a substantive dimension of public perception rather than as a secondary technical complication.

Beyond text, user-generated images provide another lens on public perception by showing what people choose to photograph and how they compose heritage scenes. The idea of a “tourist gaze” highlights that tourist photography and attention follow recurring visual routines, where certain iconic viewpoints and motifs become highly reproduced (Urry, 1992). At the same time, images can capture experiential and material details that may not be fully expressed in writing. For heritage places, this visual layer is particularly important, because visitors often respond to architectural forms, streetscapes, interiors, and symbolic landmarks in ways that are directly visible in photographs.

These developments are highly relevant for World Heritage research and management. UNESCO World Heritage Sites are both protected places and socially contested landscapes, where perception matters for interpretation, legitimacy, and sus-

tainable tourism governance. UNESCO policy documents increasingly stress the importance of monitoring change, managing tourism pressure, and engaging communities and stakeholders in stewardship (UNESCO et al., 2013; UNESCO World Heritage Centre, 2024). At the same time, critical heritage scholarship argues that official narratives often privilege monumental architecture and expert-defined values, while everyday meanings and community voices can remain less visible (Smith, 2006; Labadi, 2013). This creates a practical and conceptual need for approaches that make public narratives more observable, while still recognising their platform- and language-specific biases.

Against this background, analysing multilingual UGC across multiple platforms offers a promising way to examine how places are represented in contemporary digital discourse. It enables researchers to identify recurring themes and evaluative tones, to compare how different publics frame the same place, and to explore how textual and visual representations relate to each other at scale. For urban World Heritage sites, such evidence can complement expert frameworks by revealing which aspects of heritage are emphasised, which are ignored, and how everyday visitor experiences intersect with formally stated heritage values.

1.2. Motivation

The study is motivated by conceptual and practical gaps in current research on public perception of World Heritage. A first gap concerns the types of user-generated content and platforms that are typically analysed. Many studies rely on image-sharing platforms such as Flickr or Instagram, where data mainly consist of geo-tagged photos and short tags or captions (Ginzarly et al., 2019; Bai et al., 2023). These sources are valuable for mapping where people take pictures and which landmarks receive the most visual attention, but the associated textual information is often too brief to capture how visitors interpret heritage in their own words. In contrast, travel and map platforms such as Google Maps, TripAdvisor, and Ctrip host longer reviews in which users describe experiences, justify evaluations, and comment on atmosphere, crowding, or authenticity. At the same time, these platforms also attach user photos to the same places and visits. This thesis therefore focuses on review-based platforms and treats reviews and linked images as a combined record of visitor perceptions.

A second gap relates to the relationship between public expressions and official heritage frameworks. World Heritage designation follows a largely top-down process in which experts evaluate sites against standardized criteria and formal documents

are produced to define Outstanding Universal Value (OUV). This process often emphasises tangible attributes, architectural quality, and conservation-oriented values (Cunha Ferreira et al., 2024). Meanwhile, residents and tourists develop their own understandings through lived experience and everyday use of space, which may place more weight on social, emotional, and practical dimensions (Johnston, 2017). Although these community perceptions are central to how heritage is actually encountered, only a limited number of studies have systematically compared user-generated narratives with formally stated OUV criteria to examine where public framings converge with, or diverge from, expert narratives.

A further motivation concerns analytic scope. Because World Heritage tourism is inherently multilingual and mediated through different platforms, perception signals can shift with language community and platform conditions (Xiang et al., 2017; Cenni and Goethals, 2020). Rather than treating this as a limitation, this thesis places multilingual and cross-platform comparison at the centre of the research design. Doing so supports a more differentiated reading of public perception by revealing which themes and value framings persist across languages and platforms, and which vary with digital context.

The case of Macao makes these issues especially significant. The Historic Centre is promoted in official documents as a living testimony of East-West cultural exchange, while simultaneously functioning as a dense, mixed-use urban area shaped by tourism and casino-driven development pressures (Cultural Institute of the Macao SAR Government, 2005). The Macao SAR has adopted a Cultural Heritage Protection Law and a Protection and Management Plan for the Historic Centre that set out legal and planning instruments to safeguard OUV under conditions of rapid change (Macao Special Administrative Region, 2013). In this context, understanding how different visitor groups describe and visualise the Historic Centre across languages and platforms is relevant not only for academic debates, but also for local heritage management and communication.

Against this background, the thesis develops a place-based, multi-modal framework that links multilingual reviews and user-generated images to the OUV criteria of the Historic Centre of Macao. By organising digital traces at the level of officially designated heritage components and comparing patterns across platforms and languages, the study provides a structured way to examine how everyday public narratives relate to expert-defined heritage values.

1.3. Research Questions

Building on this motivation, the thesis is structured around four research questions that move from visitors’ textual narratives, to their relationship with the Outstanding Universal Value (OUV) of the Historic Centre of Macao, to visual representations, and finally to the links between text and images. Together, these questions provide a framework for examining how public perceptions expressed in user-generated content relate to the expert-based OUV framework in a concrete World Heritage city.

1. How do visitors describe and emotionally evaluate the Historic Centre of Macao in their multilingual online reviews, in terms of recurring themes, values, and everyday experiences across heritage components, platforms, and languages?
2. To what extent, and in what ways, do these public narratives align with the Outstanding Universal Value (OUV) criteria adopted for the Historic Centre of Macao?
3. How are the heritage components of the Historic Centre of Macao visually represented in user-generated images, and what system-wide visual themes and language-sensitive preferences emerge from these photographs?
4. How do textual and visual user-generated content complement or diverge from each other in representing the heritage values of the Historic Centre of Macao at the component level?

The first question focuses on the content and tone of visitors’ multilingual reviews, treating them as narratives of everyday experience. The second question links these narratives to the four OUV criteria for Macao, asking where public expressions converge with or diverge from expert definitions. The third question turns to user-generated images and examines which visual motifs and patterns of attention dominate across components and languages. The fourth question brings these strands together by analysing how textual and visual signals intersect within the same heritage components. These research questions guide the design of the data collection, the analytical pipeline, and the structure of the thesis.

In the remainder of the thesis, the term *heritage component* is used for the 27 officially designated monuments and urban areas that together form the Historic Centre of Macao. The term *user-generated content (UGC)* refers to online reviews and user-uploaded images collected from Google Maps, TripAdvisor and Ctrip, as detailed in 3. When discussing *public perception* or *public narratives*, the focus is

on how visitors express their views in these textual and visual traces, rather than on all possible opinions held by residents or other groups.

1.4. Thesis Structure

The thesis is organised into six chapters.

Chapter 2 reviews relevant literature on public perceptions of heritage, user-generated content and social media platforms, and computational methods for analysing textual and visual data in urban heritage settings. Chapter 3 introduces the study area, the heritage components of the Historic Centre of Macao, and the review and image datasets collected from Google Maps, TripAdvisor, and Ctrip. Chapter 4 presents the analytical framework, including the multilingual sentiment and keyterm pipeline, the two-stage OUV alignment procedure, the visual theme model, and the cross-modal component profiles.

Chapter 5 reports the empirical results, starting from the anchor case of the Ruins of St Paul's and then extending to system-wide textual, visual, and cross-modal patterns across all heritage components. Chapter 6 discusses these findings in relation to the research questions, existing work on tourist gaze and digital heritage, and ongoing debates on World Heritage management and participation in Macao. Chapter 7 concludes the thesis by summarising the main contributions, reflecting on limitations, and outlining directions for future research and practice.

2. Related Work

2.1. Public Perception and Urban Heritage

2.1.1. UNESCO World Heritage and OUV

The *Operational Guidelines* (UNESCO World Heritage Centre, 2024) for the Implementation of the *Convention Concerning the Protection of the World Cultural and Natural Heritage* set the basic rules for the World Heritage system. They explain what counts as heritage and how sites are chosen. The *Convention* divides heritage into several main groups. Cultural heritage includes monuments, groups of buildings, and sites that are important for their historical, artistic, or scientific value. Natural heritage covers physical formations, geological features, and natural sites that are important for science or conservation. Some properties meet both cultural and natural definitions and are listed as mixed heritage. Cultural landscapes are another category. They show the combined work of people and nature and reflect how societies develop in relation to their environment (UNESCO, 1972).

The central idea in this system is Outstanding Universal Value (OUV). The guidelines define OUV as cultural or natural importance that is “so exceptional” that it goes beyond national borders and is meaningful for all people, now and in the future (UNESCO World Heritage Centre, 2024). This means that protecting these places is not only the duty of one state. It becomes a shared responsibility of the international community.

A site does not become World Heritage only because it is locally important. The World Heritage Committee evaluates each nominated property against strict conditions (paragraphs 77–95 in the Guidelines). First, the property must meet at least one of ten criteria. Criteria (i) to (vi) are usually used for cultural heritage, for example when a site is a masterpiece of human creative genius or shows an important cultural tradition. Criteria (vii) to (x) are used for natural heritage, such as sites with exceptional natural beauty or key ecological and biological processes (UNESCO World Heritage Centre, 2024).

However, fulfilling these criteria is not enough. The property must also meet the conditions of authenticity and integrity. Authenticity mainly concerns cultural sites. It asks whether the form, materials, design, use, and traditions that express the value are truthful and reliable. Integrity applies to both cultural and natural sites and refers to the wholeness and intactness of the property and its setting. In addition, the site must have an adequate management system to ensure long-term protection. When a property is inscribed, the Committee adopts a Statement of Outstanding Universal Value. This statement becomes the key reference document for all future protection and monitoring efforts (UNESCO World Heritage Centre, 2024).

These rules were first written for individual monuments and clearly bounded sites. Many World Heritage properties today, however, are living cities and large urban areas. To address this complexity, UNESCO adopted the *Recommendation on the Historic Urban Landscape* (HUL) in 2011 (UNESCO, 2011). This approach goes beyond saving single buildings. It looks at the whole city as a set of overlapping layers of value. Bandarin and Van Oers (2012) explain that HUL combines urban morphology, built fabric, and social practices. It allows heritage protection to be planned together with urban development. Veldpaus et al. (2013) add that this view shifts attention away from a static list of objects. Instead, the city is treated as a dynamic landscape that changes with the daily life of its residents.

For this thesis, OUV and HUL provide the formal and conceptual background. OUV offers a clear, expert-defined value framework. HUL stresses that these values sit in a living, changing urban environment. This supports the use of Geographic Information Retrieval (GIR) analysis later in the thesis. The aim is to examine how public narratives attached to specific places in a historic city relate to the official OUV description.

2.1.2. Heritage Interpretation and Public Perception

Heritage conservation practice has changed over the past decades. Decisions were once mainly made by experts, such as historians, architects, and planners. Today, international organizations recognize that public participation is essential. Policy documents now stress that communities should not only be informed but should also help shape how heritage is managed. UNESCO, for example, states that communities need to understand the site and encourages states to integrate heritage into everyday community life (UNESCO et al., 2013; UNESCO World Heritage Centre, 2024). This shift has both practical and ethical reasons. It responds to the idea that heritage policies work better when they reflect local values and that people have a

right to take part in decisions about places that shape their lives.

Scholars support this change by rethinking what heritage is. Smith (2020) argues that heritage is not only a group of physical objects. It is also an emotional and social process. People use heritage to build and maintain social bonds. This view connects closely to the idea of “sense of place”. According to Tuan (1977), sense of place emerges from how people experience and imagine a place in their everyday life. It includes place attachment (emotional ties) and place identity (how people see themselves through places). Sense of place is subjective and dynamic. It changes over time as people’s relationships with a place evolve. By contrast, OUV is written down in official documents and tends to be more static. This contrast highlights a key tension. Expert documents try to fix values in a stable way, while public meanings are fluid and constantly renegotiated.

Empirical work also supports this more community-centered view of heritage management. A comprehensive review by Foroughi et al. (2023a) examines over 120 studies in urban and heritage planning. The review identifies factors that influence public involvement, such as governance structures, trust, and access to information. Other studies show that higher levels of participation, such as collaboration and empowerment, help bring heritage policies closer to community values (Poulios, 2014; Li et al., 2020). ICCROM case studies document how community action can protect historic buildings even when these buildings lack formal national recognition (Katrakazis et al., 2018). These examples show that formal policy is not enough. Heritage outcomes depend on how people feel about places and how they act on these feelings in daily life.

These developments are important for this thesis. They suggest that heritage management should pay attention to public perception, not just expert opinion. Today, many of these perceptions appear in digital form, for example in online reviews and shared photos. These digital traces can be read as everyday interpretations of heritage sites.

2.1.3. Expert Narratives vs. Public Narratives

Even though heritage policies increasingly emphasize participation, many researchers argue that the World Heritage system still largely operates through an expert-driven logic. In practice, OUV criteria often support what has been described as an Authorized Heritage Discourse (AHD) (Smith, 2006). This discourse privileges expert knowledge, national institutions, and monumental values, while giving limited space to everyday uses and lived meanings. Nomination dossiers, management plans, and

periodic reports play a central role in formalizing this discourse by defining why a site is significant and how it should be protected. However, these documents tend to marginalize social, emotional, and experiential dimensions that local communities and visitors associate with heritage sites (Labadi, 2013).

The dominance of expert narratives is further reinforced by the political nature of World Heritage decision-making. As shown by Meskell (2018), evaluations and Committee decisions are not purely technical processes but are shaped by diplomacy, national interests, and broader geopolitical contexts. Similarly, Harrison et al. (2020) argues that heritage should not be understood solely as the preservation of material traces of the past, but as an ongoing process that actively shapes present and future worlds. These critiques suggest that official OUV descriptions are necessarily selective: they foreground certain values while downplaying ethical, emotional, and everyday aspects that may be meaningful to different groups.

As a result, a structural gap often emerges between expert narratives and public narratives of heritage. Empirical studies show that citizens and visitors frequently value qualities such as atmosphere, daily use, or visual enjoyment (Ginzarly et al., 2019), which remain marginal in the technical language of OUV criteria. Consequently, the official story articulated in nomination and management documents does not always align with how heritage sites are experienced on the ground. This gap highlights the need for bottom-up evidence that captures how people actually talk about, use, and make sense of heritage in everyday contexts.

Importantly, this need does not disappear once a site has been formally designated as World Heritage. While OUV statements establish a site’s official significance at a particular moment, heritage meanings are not fixed by designation alone. Heritage has long been understood as a dynamic and socially constructed process rather than a stable set of attributes (Smith, 2006; Graham et al., 2000). As sites continue to be lived in, visited, photographed, and discussed, ongoing interactions actively shape how heritage is understood and valued over time (Waterton and Smith, 2010). Public narratives therefore play a crucial role in how heritage is encountered in practice, extending beyond the relatively static definitions articulated in official documents.

From a governance perspective, attending to public narratives is particularly relevant in contexts where heritage management seeks to balance conservation objectives with tourism, local use, and community engagement. Values-based approaches to heritage management emphasize that long-term sustainability depends not only on expert assessments, but also on social acceptance and perceived legitimacy (Pendlebury, 2013). If official narratives diverge too strongly from public perceptions and experiences, management strategies risk overlooking experiential dimensions that

matter to users or losing social legitimacy. In this sense, public narratives can be understood as a form of informal, large-scale feedback that reveals which heritage values resonate with visitors and residents, which remain marginal and generate appreciation or tension over time (Labadi and Logan, 2015).

In this context, digital platforms offer new opportunities to study public narratives systematically. Online reviews and photographs shared on platforms such as Google Maps, TripAdvisor, and Ctrip document how heritage sites are described, evaluated, and visually represented by the public at scale. Because these data are often spatially referenced, they are particularly well suited for spatial analysis. Analyzing such data makes it possible to examine how public narratives relate to, reinforce, or diverge from officially defined OUV frameworks in contemporary heritage contexts.

2.2. Tourism, Social Media Platforms and UGC

2.2.1. UGC in Tourism and Destination Image

Tourism studies describe public perception through the concept of Tourism Destination Image (TDI). TDI refers to how people see and feel about a place as a travel destination. To analyze this, researchers usually split the image into three parts. Baloglu and McCleary (1999) describe a cognitive component, which refers to people's knowledge and beliefs about a place (for example, "clean streets" or "historic buildings"), and an affective component, which covers emotions and feelings toward that place (for example, "pleasant" or "boring"). Later, a conative component was added to capture actions or intentions that follow from these thoughts and feelings, such as the choice to visit or recommend the destination (Agapito et al., 2013).

Tourism does not take place in an abstract space. It happens in specific streets, squares, and buildings. Therefore, many authors stress that destination image also has a spatial side. Classic work by Lynch (1960) and Pocock and Hudson (1978) shows that recognizing a city involves both identifying elements (such as landmarks and paths) and understanding their spatial arrangement. The relation between "what" is seen and "where" it is located forms a mental map that guides movement and behavior. This perspective is important for heritage studies. It suggests that a historic centre should be studied not as a single unit, but as a network of geographically distinct attractions that together form the destination image.

Tourists engage with both the psychological and spatial dimensions mainly through visual consumption. Urry (1992) introduced the idea of the Tourist Gaze. This

concept describes how visitors search for and consume particular views that match their expectations of what is special or “worth seeing”. In the past, this gaze was expressed through personal photography. Today, it is widely expressed on digital platforms, where visitors take photos and share them online.

This process reveals possible conflicts between projected images and perceived images. Place marketing studies distinguish between projected images, which are created by official bodies such as tourism boards (Kotler, 2002), and perceived images, which are the mental images formed by visitors (Bramwell and Rawding, 1996). Marine-Roig and Ferrer-Rosell (2018) argue that official images are often idealized and carefully staged. User images, by contrast, may show the more “messy” reality of crowding, noise, or unexpected uses of space.

Until recently, surveys were the main way to measure these differences. However, surveys have clear limits in size, cost, and frequency. There has therefore been a strong shift toward using User-Generated Content in tourism research. Chen and Law (2016) argue that UGC is a powerful data source because it is unsolicited, cheaper to obtain, and often more honest than responses to a questionnaire. Stepchenkova and Zhan (2013) show that online photos and maps can effectively capture the perceived image of a country and reveal how tourists move through and frame that space. These works suggest that digital footprints are useful for studying both what people notice and where they go.

Online content also shapes future travel decisions. Potential visitors see images and reviews, build expectations, and then try to reproduce similar scenes when they visit. This feedback loop has been described as a hermeneutic circle of representation. Halegoua (2020) uses the term digital placemaking to show that digital media does not just represent places. It actively shapes how places are experienced and valued. Boy and Uitermark (2017) demonstrate how specific spots become “Instagrammable moments” and attract concentrated attention. This selective visualization can change how crowds move through heritage sites. It may draw visitors toward a few iconic viewpoints and away from less visible but still important locations.

These ideas support the use of UGC to study how visitors see and feel specific sights in a historic centre and how these images and texts might reinforce or challenge the official OUV narrative.

2.2.2. Platforms as Different Tourist Cultures

Social media platforms are not neutral channels. Their design shapes how people write, photograph, and share content. Researchers describe these design features as platform affordances (Bucher et al., 2018). Technical elements such as character limits, photo layouts, and mobile interfaces influence what users post. As a result, reviews of the same place can look and feel very different across platforms.

TripAdvisor is linked to a specific travel advice culture. Tourists often use it for pre-trip planning. Reviews on TripAdvisor therefore tend to be longer and more detailed. Xiang et al. (2017) find that TripAdvisor reviews are often of higher perceived quality and more useful for planning than reviews on some other platforms. For this reason, TripAdvisor has long served as a standard data source in tourism research (Chua and Banerjee, 2013; Jeacle and Carter, 2011). However, it also has limits for heritage interpretation. Koufodontis and Gaki (2022) analyze reviews in 105 World Heritage cities and show that tourists are not always aware of the UNESCO status of the places they visit. Awareness varies with micro-level factors and with how attractions are presented on social media.

Google Maps represents a more geo-anchored and mobile-first culture. Most users access it on their phones while they are on the move. They often write short reviews directly after a visit. These reviews are usually brief and informal but are tightly linked to specific points on the map. This has two effects. First, the platform accumulates very large numbers of reviews, which attracts researchers who need big datasets. Second, the strong link between review and location makes Google Maps particularly suitable for GIR analysis. Recent work shows that, for the same destination, Google Maps often yields far more reviews than TripAdvisor (Haris et al., 2023; Alam et al., 2024). Mellinas and Sicilia (2024) add that Google reviews tend to focus on immediate, on-site experiences.

Ctrip reflects a distinct Chinese tourist culture. It differs from Western platforms in both language style and visual focus. Yang (2025) argue that Ctrip users aim to build friendly relationships and avoid too strong self-assertion. They use hedges and polite expressions more often than “I”-statements that claim authority. Strong personal voice is more common in complaints. Huai and Van de Voorde (2022) compare reviews from Shanghai on Ctrip with reviews from Brussels on Western platforms. They find that Ctrip reviews are more sensitive to visual environmental attributes such as color and water, while Western reviews focus more on experiential qualities like calmness. This suggests that Ctrip expresses a specific gaze that is socially modest and strongly visual.

Each platform thus captures only a part of the tourist market and a specific discursive culture (Xiang et al., 2017). Relying on one platform alone can produce a biased picture of public perception. Therefore this study will compare Google Maps, TripAdvisor, and Ctrip. The aim is to build a more complete view of how different cultural and linguistic groups talk about and visually represent the same heritage sights.

2.2.3. Heritage-Specific Uses of Social Media

Social media data are now widely used to study historic cities and cultural sites. This work is sometimes described under the label “Smart Heritage”. Early studies often used such data for general tourism indicators, such as visitor satisfaction or volume. For example, Santa-Cruz and López-Guzmán (2017) analyzed visitors in Córdoba to measure their satisfaction with the urban heritage setting.

More recent studies look more closely at specific heritage values. Ginzarly et al. (2019) use geotagged Flickr images to explore how people perceive values in urban heritage areas. Their results show that citizens are often drawn to less tangible aspects, such as atmosphere and everyday use. These aspects do not always receive the same attention in expert-led discussions.

New computational methods now try to link digital data directly to formal heritage frameworks. Bai et al. (2021) develop a deep learning model that can classify text according to UNESCO OUV-related concepts. This shows that artificial intelligence can be trained on expert language and then used to detect similar patterns in other texts. Bai et al. (2023) apply AI methods to urban heritage images in Rome. They treat social media photos as geo-anchored signals and map how people visually consume different parts of the city. Together, these studies show that it is possible to connect OUV frameworks to large digital datasets of text and images.

Other work focuses on specific architectural features. Foroughi et al. (2023b) use Instagram and Twitter to study the windcatchers of Yazd, a key OUV attribute. They compare how residents and tourists present these structures and find both shared and divergent visual focuses.

Despite these advances, a major gap remains. Few studies use a systematic framework to measure how strongly public UGC aligns with particular OUV criteria (such as criterion ii versus criterion iv). Most work does not yet quantify this alignment in a way that allows direct comparison across all sites in one city. This study will address this gap by building a GIR-based framework. The framework links text and

images from multiple platforms to the official OUV claims and produces indicators of alignment.

2.3. Language and Multilingual Public Perception

2.3.1. Language as a Lens of Culture and Heritage Meanings

In addition to spatial patterns, the linguistic diversity of UGC plays a crucial role in how heritage is perceived. Language is more than a neutral tool for communication. It is a lens through which people interpret the world. It shapes what they notice and how they describe it. He (2010) argues that a heritage language is deeply sociocultural. It is tied to the group that speaks it. Language becomes a way to identify with that group and to carry cultural values. When tourists write reviews in different languages, they are therefore not simply translating a fixed content. They express their cultural identities and value systems through that language.

In global heritage cities, multiple languages coexist in streets, shops, and public signs. Scholars describe this as a linguistic landscape. Gorter (2006) suggests that such landscapes are not random. They form a systematic symbolic construction of public space. The presence or absence of certain languages on signs sends messages about identity and power. In later work, Gorter (2013) argues that multilingual cityscapes can be read as a record of social change over time. They show shifts in migration, economics, and policy.

In the digital age, these patterns extend into online space. The set of languages used in reviews and posts around a site can be seen as an online linguistic landscape. The visibility of different languages in UGC signals which groups are more active and what kinds of meanings they attach to the site.

2.3.2. Multilingual Social Media in Tourism

Social media enhances these dynamics through a process known as mediatization. Agha (2011) notes that communication links individuals to wider institutions and social processes. Online platforms turn these links into enduring records. They also host metalanguage, which means talk about language itself. Jaworski et al. (2012) show that such comments can reveal deep beliefs about what counts as “correct” or “authentic” language. Blommaert et al. (2009) describes language policing in online spaces, where users correct or criticize each other’s language use. These practices

show that language on social media is closely connected to questions of power and identity.

Empirical studies confirm that tourists' languages shape their evaluations. Hale (2016) analyze TripAdvisor ratings for London attractions in 16 languages and find clear differences in average scores between language groups. Liu et al. (2017) compare Chinese reviews with reviews in eight other languages and examine attributes such as room quality, cleanliness, and service. They show that different language groups assign different importance to these attributes.

Language also affects how dissatisfaction is expressed. Goethals (2016) study comments in German, French, and Spanish on Booking.com and find distinct patterns of judgment. Cenni and Goethals (2017) look at negative TripAdvisor reviews in English, Dutch, and Italian and analyze the speech acts used in complaints. They show that language groups differ in strategies such as exaggeration or softening. Taken together, these studies show that rating patterns, highlighted attributes, and narrative styles all vary with the language of the review.

At the same time, language does not map neatly onto nationality or culture. Liu et al. (2017) note that English often functions as a shared second language in tourism. Many English-language reviews are written by non-native speakers from diverse backgrounds. This means that language groups are not the same as national groups. Therefore, this analysis focuses on how reviews in different languages talk about heritage sights and OUV-related themes, without assuming simple one-to-one links between language and nationality.

2.3.3. Translation and Cross-Language Comparison

Comparing content across languages creates methodological challenges. One key challenge is conceptual equivalence. Translation is more than replacing words. It involves transferring cultural meanings. Temple and Young (2004) argues that translators are not neutral tools. They rewrite texts based on their own understanding. This can introduce bias, especially when concepts do not fully match between cultural contexts. Choi et al. (2012) warn that word-for-word translation often fails to capture the full sense of the original. Kirilenko et al. (2018) note that standard dictionaries and generic tools may miss tourism-specific nuances, leading to semantic mismatches.

Many recent studies rely on machine translation to handle large volumes of text. This approach is efficient but has limits. Vanmassenhove et al. (2019) show that ma-

chine translation systems tend to reduce lexical diversity. Human writers use many different words and expressions, while machine output often prefers frequent, safe options. This can flatten emotional nuance and reduce variation in style. In tourism reviews, this may mean that subtle cultural differences are lost after translation.

To reduce these problems, researchers sometimes use back-translation. A text is translated into a target language and then translated back to the original language to check if meaning is preserved. Choi et al. (2012) point out that back-translation alone is not enough. The translator also needs deep cultural knowledge to understand context. Otherwise, they may correct words but still miss underlying meanings. Kirilenko et al. (2018) conclude that for complex or noisy datasets, fully automated analysis remains weaker than careful human processing.

At the same time, it is not realistic to manually translate and code every review in a large UGC dataset. A practical compromise is needed. In this study, automated translation to English is used to create a common text layer for embedding and classification. Original languages are kept for language-specific word clouds and for interpreting cross-language differences. A sample of translated texts will be checked manually to understand common translation errors and their likely effects on the analysis. This mixed strategy follows the literature’s warning about bias, while still making large-scale, multilingual comparison feasible.

2.4. NLP and Computer Vision for Tourism/Heritage

2.4.1. Text Analytics in Tourism and Heritage

The rise of social media has produced large amounts of text about tourism and heritage. Researchers now use NLP to handle this volume and to study how people feel about places and what they talk about most. In this context, two key tools are sentiment analysis and keyterm extraction.

Sentiment analysis aims to detect the emotional tone of a text. Early work often used sentiment lexicons and simple rules. For example, Clarizia et al. (2018) combined a lexicon-based sentiment method with basic topic grouping to evaluate public reactions to cultural heritage events. Their results showed that such methods can separate broadly positive and negative views, but they also suggest that fine-grained or mixed emotions are harder to capture.

More recent studies increasingly rely on machine learning and deep learning for

sentiment analysis. Catelli et al. (2023) used BERT- and ELECTRA-based models to detect fake reviews in the Italian cultural heritage domain. They built a new dataset and showed that deep models can capture complex patterns of language that simple lexicon methods miss. Their work also used sentiment scores as an extra signal for detecting misleading or suspicious content, which shows that sentiment analysis can support more advanced diagnostic tasks beyond simple polarity.

Keyterm extraction is another important tool for text analytics. Instead of grouping words into hidden topics, keyterm methods identify the most distinctive words and phrases in a corpus or subset of texts. Classical approaches rely on frequency and TF-IDF scores. More recent unsupervised methods, such as RAKE and YAKE, use local co-occurrence, position and other simple features to rank candidate terms without supervision (Rose et al., 2010; Campos et al., 2020). These methods are useful because they produce concrete, interpretable terms that can be linked back to specific places, languages, or platforms.

In tourism and heritage research, many studies also rely on non-parametric statistics when data are ordinal or not normally distributed. For example, Adiwibowo et al. (2015) used Spearman rank correlation to test how public appreciation of historic buildings relates to intention to visit reused heritage buildings in Bandung. Rafael and Almeida (2017) applied Kolmogorov–Smirnov tests to examine how different socio-demographic groups perceive destination image in an online setting. In cultural ecosystem services studies based on social media, Oteros-Rozas et al. (2018); Chen et al. (2024) used Spearman correlations to link social media indicators of cultural ecosystem services with landscape features and park characteristics. These examples support the use of rank-based methods, such as Spearman correlation and Wilcoxon-type tests, to compare keyword rankings and other derived indicators across languages, platforms, and sites in this thesis.

Some heritage studies have also used topic models to uncover value themes, for example in historic districts or local narratives (Pan et al., 2024; Tenzer and Schofield, 2024). These works show that automated methods can reveal recurring themes related to heritage, leisure, or spiritual experience. However, topic models are harder to interpret and harder to compare across languages, and they often require substantial manual reading and relabelling.

This study therefore focuses on sentiment analysis and keyterm extraction rather than full topic modelling. Sentiment analysis is used to measure how visitors feel about the Historic Centre of Macao and its sites. Keyterm extraction is used to identify significant words and phrases in different languages, platforms, and site groups. These keyterms are then used to build word clouds, ranked lists, and non-

parametric statistical comparisons across languages and platforms. This approach keeps the methods transparent and closely tied to the actual words used by visitors, while still making it possible to analyse patterns at the scale of the whole World Heritage area.

2.4.2. Sentence Embeddings and Claim Alignment

Many recent NLP methods work not at the level of single words, but at the level of sentences. Sentence embeddings map each sentence to a point in a high-dimensional space. In this space, semantically similar sentences lie close to each other. Models such as Sentence-BERT build these embeddings by fine-tuning BERT with a contrastive loss on sentence pairs (Reimers and Gurevych, 2019a). This makes cosine similarity a useful way to compare meaning across many sentences.

Embedding methods are useful when we want to link public texts to formal heritage concepts. Bai et al. (2021) trained deep learning models on UNESCO statements of Outstanding Universal Value (OUV). Their WHOSe Heritage framework treats OUV criteria as labels and learns how they are expressed in expert documents. This work shows that OUV can be turned into computable categories and that soft labels can capture overlaps between criteria.

For this thesis, the same idea can be adapted at a different level. Besides classifying the reviews into general UNESCO statements of OUV by Bai et al.’s (2021) model, the goal is also to compare them with fine-grained OUV claims from the Macao nomination file. Sentence embeddings make this possible. Each review sentence and each OUV claim is embedded in the same vector space. A similarity score then indicates how closely a visitor’s wording matches an official claim. This supports the construction of claim–sentence alignment tables, which link everyday language back to formal OUV categories.

2.4.3. Vision Methods in Tourism and Heritage

User-generated images offer another way to study how people experience heritage. Computer vision methods can process these images at scale and detect patterns that are not visible through manual inspection alone.

Early visual studies often relied on basic statistics, such as where photos were taken or how many photos existed per site. Recent work uses deep learning to go further. Payntar (2021) combined geotagged photos and machine learning to study tourist

movement and visual preferences at archaeological sites around Cuzco, Peru. They showed how clustered visual scenes, such as mountains, stonework, or terraces, align with different parts of the heritage landscape.

Other studies use semantic segmentation or scene parsing to understand what appears in each image. Zhao et al. (2025) used SegFormer-based models and sentiment analysis on social media photos to assess aesthetic and emotional values in landscapes. Their work illustrates how deep models can extract rich visual features (for example, sky, water, vegetation, buildings) and tie them to people’s reactions.

In tourism and heritage research, CNN-based features are now used to cluster photos and study how people look at places. Bai et al. (2022a) build the Heri-Graphs dataset, where they treat social media posts as multi-modal nodes and use image features together with text and location to group patterns of heritage values across cities. In a more detailed city case, Bai et al. (2023) apply VGG-based features, dimensionality reduction and clustering to social media images from the Testaccio district in Rome, and then interpret each cluster by reading sample photos to identify recurring visual themes. Calcagni et al. (2022) combine Flickr photos and user-generated tags to build an empirically grounded typology of relational values in a peri-urban park, using careful manual coding of both images and text. Taken together, these studies show that visual clusters often reflect distinct ways of “seeing” a site and can be read alongside textual descriptions. This dual view is central to this thesis, which treats geo-tagged photos as signals of tourist gaze that complement narratives in online reviews.

2.4.4. Strengths and Pitfalls of NLP and Vision Methods

NLP and computer vision methods offer clear advantages for heritage research. They allow researchers to process very large datasets, detect hidden patterns, and link public narratives to formal value frameworks such as OUV (Bai et al., 2021; Tenzer and Schofield, 2024). They also make it possible to work at multiple scales, from global comparisons of OUVs to local analyses of specific districts or sites.

However, these methods also have important limits. Topic models and classifiers can misread irony, local idioms, or mixed feelings. Deep models such as BERT are powerful but can behave as “black boxes,” making it hard to explain why a specific sentence was linked to a certain value. Domain shift is another problem: models trained on general corpora may not handle heritage-specific terms or multilingual reviews very well. Similar issues appear in vision. Models trained on generic image datasets may misclassify heritage scenes or focus on irrelevant features when applied

to historic cities (Zhao et al., 2025).

Because of these issues, many studies stress the need for human-in-the-loop validation. Tenzer and Schofield (2024) combine topic modelling with manual interpretation to refine value categories and avoid misleading themes. Bai et al. (2021) evaluate their OUV classifier with expert judgement to check that predictions align with heritage practice. Visual studies such as Payntar (2021) also combine automatic analysis with qualitative reading of images.

This thesis follows the same principle. Automated methods are used to scale up analysis, but manual checks and an anchor case (Ruins of Saint Paul’s) are used to calibrate thresholds, tune labels, and document errors. In this way, the technical pipeline supports, rather than replaces, expert and community-based understanding of heritage.

3. Materials

3.1. Overview

This chapter introduces the materials used in the analysis and explains how they were assembled into a structured dataset for the study. The materials include user-generated reviews in multiple languages, user-uploaded images linked to the officially designated heritage sites, and documentary sources that describe the Outstanding Universal Value (OUV) of Historic Centre of Macao. Together, these sources provide complementary perspectives on how visitors encounter, interpret, and evaluate the heritage environment.

3.2. Case Study: Macao World Heritage

The Historic Centre of Macao is selected as the case study for this research. The core area covers approximately 16.2 hectares and was inscribed on the UNESCO World Heritage List in 2005 under the title *Historic Centre of Macao*¹. The property is officially recognised for its Outstanding Universal Value (OUV) as a unique example of sustained cultural exchange between East and West (Cultural Institute of the Macao SAR Government, 2005). According to the decision of the World Heritage Committee, the site is inscribed under cultural criteria (ii), (iii), (iv), and (vi). Although the earlier nomination dossier discussed aspects related to criterion (v), this criterion was not included in the final inscription. The analysis in this thesis therefore follows the officially adopted criteria.

Although UNESCO's official descriptions of criteria (ii), (iii), (iv), and (vi) partly overlap, this thesis adopts a set of dominant analytical lenses to distinguish them in practice. Criterion (ii) is interpreted as focusing on mechanisms of cultural and intellectual exchange; criterion (iii) on Macao as a long-term contact zone and space of

¹ <https://whc.unesco.org/en/list/1110/>, accessed 10 April 2025. Licensed under CC-BY-SA IGO 3.0.

cultural fusion; criterion (iv) on the material urban fabric and architectural ensembles that embody these encounters; and criterion (vi) on the circulation of ideas and their links to wider socio-political transformations in China. Together, these criteria define the heritage significance against which contemporary public perspectives can be interpreted. In short:

- (ii) cultural–intellectual exchange mechanisms
- (iii) contact zone & cultural fusion
- (iv) material urban fabric & architectural ensembles
- (vi) circulation of ideas & socio-political change

The Historic Centre has developed into a complex urban landscape shaped by more than four centuries of interaction between Chinese and Portuguese communities. Macao remained under Portuguese administration from the mid-16th century until its handover to China in 1999. During this period, the city functioned not only as a Portuguese colony but also as an important regional trading port used by merchants from various European countries. Its position within wider maritime networks brought a steady flow of goods, people, and ideas, further enriching the cultural layers that characterise the historic urban landscape today.

Visitors all over the world share their experiences of the Historic Centre in a range of languages. This study focuses on English, traditional Chinese, and simplified Chinese reviews, which together represent key segments of the public engaging with Macao’s heritage sites.

The study area includes a set of officially designated heritage sites that reflect the city’s historical depth and cultural symbolism. Notable examples include the Ruins of St Paul’s, originally part of a 17th-century Jesuit complex, and the A-Ma Temple, established in 1488. These and other religious, civic, and public spaces illustrate the diverse functions of the Historic Centre and its continuing role within Macao’s contemporary urban fabric.

In this thesis, the basic spatial unit of analysis is the heritage component, corresponding to one of the World Heritage elements that make up the Historic Centre of Macao. Each component is represented by a unique identifier (`poi_id`) and associated with a geographic location (centroid coordinates) derived from official inventories and mapping sources. The `poi_id` key links all textual reviews and user-generated images to their physical location and thus connects digital user narratives to the underlying spatial structure of the site. Throughout the thesis, the term *heritage component* (also referred to as *component* or *POI*) is used to denote one of

the individual elements that together form the World Heritage property.

3.3. Textual User-Generated Content Collection

In this thesis, textual user-generated content (UGC) is operationalized as online reviews posted by users on digital platforms. Data are collected from a set of culturally significant locations situated within the Historic Centre of Macao, which covers 16.2 hectares. The UNESCO property comprises 30 components. Three had fewer than ten available reviews across all platforms and were therefore excluded from subsequent analysis, leaving 27 components in the analytical dataset (Figure 3.1).

To capture a broad range of visitor perspectives, reviews were collected from three travel platforms: Google Maps, TripAdvisor, and Ctrip. These platforms differ in their geographic reach, user demographics, and language distributions, and they provide a complementary set of observations on visitor experiences in the Historic Centre of Macao. In this study, attention is directed to reviews written in English, traditional Chinese, and simplified Chinese, as these languages represent key visitor groups and allow for consistent manual verification during later stages of analysis.

Across the three platforms, the combined review corpus reflects both international and regional visitor profiles. Google Maps and TripAdvisor primarily capture the experiences of travellers from around the world, while Ctrip provides substantial representation of visitors from mainland China. Together, these sources offer a multilingual and multi-sourced perspective on how different publics engage with Macao's World Heritage sites. The following subsections describe the characteristics of each platform and the procedures used for review retrieval.

3.3.1. Google Maps

Google Maps provides a substantial volume of user-generated reviews contributed by visitors from around the world. Since Google services are not legally accessible in mainland China, the platform primarily reflects perspectives from international travellers, including those from Hong Kong, Macao, Taiwan, and overseas Chinese communities. Using the open-source tool *google-reviews-scraper-pro*, reviews and images were extracted for each of the 27 officially designated heritage components within the protected area. For each site, up to the top 100 most relevant reviews per available language were retrieved, together with metadata such as rating, timestamp,



Figure 3.1.: Map of the designated heritage components. Source: (Macao Government Tourism Office, n.d.).

and user identifier.

Google Maps therefore becomes an essential source of international visitor perspectives, offering a multilingual foundation for later stages of textual analysis.

3.3.2. TripAdvisor

TripAdvisor provides an additional source of user-generated reviews from an international audience. The platform is widely used by independent tourists, and its reviews often contain reflective assessments of heritage quality, cultural interpretation, and overall visitor experience. For this study, only English-language reviews were collected.

Data extraction was carried out using the browser extension *Tripadvisor Reviews Scraper*, which enables automated retrieval of publicly available comments from individual heritage component pages. For each site, the most relevant English reviews were collected together with associated metadata such as rating, timestamp, and user identifier. These TripAdvisor reviews complement the perspectives captured from Google Maps by providing additional accounts from international visitors who may engage more deeply with cultural and historical aspects of the site.

3.3.3. Ctrip

Ctrip is one of the most widely used travel platforms in China mainland and therefore provides an essential source of simplified Chinese reviews. Comments on Ctrip often combine practical travel information with personal impressions, emotional evaluations, and reflections on cultural experiences. As such, they offer a valuable complement to the more internationally oriented perspectives captured through Google Maps and TripAdvisor.

All review data from Ctrip are written in simplified Chinese, which corresponds to the platform’s primary user base. Data extraction was conducted using *Octoparse*, a web-scraping tool that supports automated retrieval of publicly accessible content from structured web pages. For each heritage component, the most relevant reviews were collected together with metadata such as rating, and user identifier. These Ctrip reviews ensure that the experiences and viewpoints of mainland Chinese travellers are fully represented in the multilingual corpus used in this study.

3.4. Visual User-Generated Content Collection

Visual user-generated content (UGC) refers to user-shared images associated with heritage sites on online platforms. In addition to textual reviews, images provide a complementary perspective on how visitors visually engage with the heritage landscape of Macao. They often capture architectural details, urban layouts, and experiential moments that may not be fully expressed in written form. Collecting these images therefore enriches the evidence available for understanding visitor perceptions of the heritage components.

User-generated images were collected from Google Maps as part of the review retrieval process. The image collection was restricted to photographs attached directly to the reviews included in this study. For each heritage component, up to the top 100 reviews per language (English, traditional Chinese, and simplified Chinese) were retrieved, and each review contributed a maximum of four images. This constraint ensures balanced representation across users and maintains a clear association between textual content and visual material.

Images were downloaded using the same open-source tool employed for text collection, *google-reviews-scraper-pro*. Alongside the image files, available metadata such as upload timestamp and user identifier were retained when present. Across the heritage components, the number of collected images varies considerably, ranging from

around 70 photographs for sites with fewer reviews to more than 700 for the most frequently reviewed locations. Although the number of images varies by component, the combined collection offers a coherent set of visual materials that complements the multilingual review corpus.

3.5. UNESCO Documentation and OUV Knowledge Base

This study draws on a set of UNESCO and advisory body publications that clarify how Outstanding Universal Value (OUV) is defined, interpreted, and managed within the World Heritage system. At a general level, two documents provide broader context for heritage management practices: the resource manual *Managing Cultural World Heritage* (UNESCO et al., 2013) and the World Heritage Papers volume *Engaging Local Communities in Stewardship of World Heritage* (UNESCO World Heritage Centre, 2014). These publications outline practical considerations such as management frameworks, stakeholder participation, and community engagement.

The primary authoritative source for the conceptualisation of OUV is the *Operational Guidelines for the Implementation of the World Heritage Convention* (UNESCO World Heritage Centre, 2024). The Guidelines set out the formal procedures for inscription on the World Heritage List, including the cultural and natural criteria, the conditions of authenticity and integrity, and the requirements for long-term protection and management. They also provide the precise operational definition of OUV that underpins evaluations by the World Heritage Committee and its advisory bodies. Beyond offering conceptual background, these general policy documents also serve an analytical purpose in this thesis. Bai et al. (2021) note that such UNESCO descriptions were used in pre-training their WHOSe Heritage classification model, which is applied in the first stage of this study. These documents therefore provide an interpretive reference for understanding the model’s outputs and for conducting the sampling-based validation of its predictions.

In addition to these general policy texts, the construction of the OUV knowledge base for this thesis relies on a set of site-specific documents relating to the Historic Centre of Macao. Three sources are central to understanding how the property’s OUV is formally articulated. The nomination dossier submitted by China for inscription on the World Heritage List (State Party of the People’s Republic of China, 2005) presents the State Party’s proposed argument for the property’s cultural sig-

nificance and initially discusses Macao’s heritage in relation to cultural criteria (ii), (iii), (iv), and (v). In contrast, the Advisory Body Evaluation prepared by ICOMOS (ICOMOS, 2005) revises this assessment and recommends inscription under criteria (ii), (iii), (iv), and (vi), the formulation subsequently adopted by the World Heritage Committee. The extraction of Macao-specific OUV statements used in this thesis draws primarily on the ICOMOS evaluation and the official UNESCO property description available on the World Heritage Centre website (UNESCO World Heritage Centre, 2005). A full list of the extracted OUV claim-level statements is provided in Appendix A.3.

The UNESCO publication *The Historic Centre of Macao* (Cultural Institute of the Macao SAR Government, 2005) consolidates the Committee’s decision and provides the authoritative description of the property, a component-by-component presentation of the thirty heritage components, and maps illustrating the core and buffer zones. Further contextual information on components is taken from material produced by the Macao Government Tourism Office, in particular the brochure *The Historic Centre of Macao* (Macao Government Tourism Office, n.d.), which provides concise descriptions of each component alongside visitor-oriented maps. Together, these site-level documents offer complementary perspectives.

Building on these sources, the OUV-related text was reorganised into a set of concise, claim-level statements. Each claim captures a specific aspect of the heritage significance identified in the documents and is linked to one of the four adopted cultural criteria. The resulting OUV knowledge base provides a structured representation of how the property’s Outstanding Universal Value is officially framed, forming the reference framework through which the analytical chapters interpret visitor perceptions expressed in reviews and images.

3.6. Data Cleaning and Preprocessing

This section describes the preparation of textual reviews, metadata, and images prior to the analytical stages of the thesis. Since the dataset is assembled from multiple platforms, languages, and media types, a structured and reproducible preprocessing workflow is required to ensure consistency across sources. The procedures outlined here aim to normalise the data, remove noise, and create a unified dataset that can be reliably used for subsequent OUV classification and interpretation.

3.6.1. Textual Data

The review texts collected from Google Maps, TripAdvisor, and Ctrip. The raw data contain inconsistencies related to platform-specific formatting, multilingual content, and user-generated noise. The preprocessing steps followed in this study include:

- **Language identification and filtering.** Although each platform predominantly provides reviews in a specific target language (English, simplified Chinese, or traditional Chinese), automated language identification was performed to ensure that all reviews were correctly classified and that occasional out-of-language entries were excluded.
- **Removal of platform objects.** Review texts often include emojis, URLs, markup symbols, or duplicated auto-generated phrases (e.g. signatures or device labels). These elements were removed to avoid interference with downstream NLP models.
- **Sentence segmentation.** To support the sentence-level OUV classification pipeline, each review was segmented into sentences using language-appropriate rules. English reviews were processed using spaCy, while Chinese reviews were segmented using jieba and a customised rule-based splitter that removes numbering patterns, punctuation residues, and other artefacts that commonly appear in user-generated Chinese text.
- **Handling of machine-translated text.** To enable cross-language comparison within the pretrained OUV classification models, Chinese reviews were then translated into English using Google Translate. Several open-source translation models were tested during preliminary experiments. However, their output quality was insufficient for heritage-related terminology. Google Translate was therefore selected as the most reliable option for producing consistent, high-quality translations. Because machine translation can occasionally merge or split clauses, a secondary segmentation step was applied after translation to ensure that sentence boundaries remained aligned across languages.
- **Token and character level normalisation.** Unicode harmonisation, lowercasing (for English), and conversion of full-width characters (for Chinese) were applied to reduce noise and maintain compatibility with the pretrained OUV models.

3.6.2. Metadata Alignment

Since the review data originate from three platforms with different data structures and naming conventions, the primary goal of metadata alignment in this study was to harmonise the datasets into a unified format.

- **Harmonisation of platform-specific formats.** The raw datasets from the three platforms provide different fields and file formats (e.g. nested JSON from Google Maps, HTML-derived text from TripAdvisor, and Excel files from Ctrip). These were standardised into a common schema containing review ID, platform, language, component name, raw text, cleaned text, and sentence segments. All three sources were then merged into a single review-level table that forms the input to the sentence segmentation and OUV classification pipeline.
- **Consolidation of image metadata.** A parallel process was carried out for images extracted from Google Maps reviews. Each image was linked to its parent review and heritage component, and the file structures were unified across English, simplified Chinese, and traditional Chinese review sets. Only images associated with the selected top 100 reviews per language were included. The resulting consolidated image table contains file paths, component identifiers, and language labels, providing the structure required for subsequent analysis.
- **Deduplication.** Duplicate entries, which may occur due to multi-page scraping, were removed using text hashing and metadata checks.

3.6.3. Image Filtering

Images accompanying Google Maps reviews represent an important component of the dataset, offering insight into what visitors choose to photograph and emphasise.

However, raw image collections contain noise, duplicates, and irrelevant material. Image preprocessing was therefore necessary and followed these steps:

- **Extraction of images linked to top reviews.** Each Google Maps review may contain up to four images. For each language, the processing script identifies the local filenames associated with the selected top 100 reviews and extracts only the required images. Separate folders (e.g. `review_images_EN_top100`) were generated for the image sets.
- **Format and resolution standardisation.** Images were converted to a con-

sistent file format and resized to dimensions suitable for further analysis.

- **Removal of low-information and corrupted images.** Blank, dark, or corrupted files were excluded automatically.
- **Near-duplicate detection.** Perceptual hashing was used to remove near-identical images, which frequently occur when users upload multiple shots of the same scene.

All photographs linked to user reviews were first cleaned and organised into a uniform directory structure. Each of the 27 heritage components has its own subfolder (for example `a_ma_temple`, `ruins_st_pauls`) under a common root directory. Within each folder, images are stored as `.jpg` files with harmonised filenames of the form:

$$\{\text{poi_id}\}_{\text{language}}\text{-}\{\text{running index}\}.\text{jpg}.$$

A metadata table links every file to its contextual information. Each row corresponds to one image and contains at least the following fields: a unique image identifier (`image_id`, matching the filename without extension), the component identifier and name (`poi_id`, `poi_name`), the review language, the associated review identifier (`review_id`), and the local filename and path. After extracting review-linked images under these constraints and applying filtering and deduplication, the final analytical image corpus contains 14,079 photos.

3.6.4. Data Summary

A detailed overview of the number of preprocessed reviews and associated images for each heritage component is provided in Appendix A.2. Across the 27 officially recognised heritage sites, the volume of collected material varies substantially. Some components attract large numbers of visitors and therefore produce extensive multilingual review and image data, while others are less frequently visited and yield only a small number of usable entries. This variation reflects differences in public visibility and tourist mobility patterns rather than inconsistencies in data collection, and it forms an important background condition for the analyses presented in later chapters.

3.7. Legal Considerations and Limitations

3.7.1. Ethical and Legal Considerations

All data used in this study were collected exclusively from publicly accessible platforms—Google Maps, TripAdvisor, and Ctrip—where users voluntarily publish reviews and images under the terms of service of each platform. No personal data beyond what is publicly displayed were collected, stored, or analysed. Usernames were not used in any part of the analysis, and all processing steps avoided any form of individual-level identification. The study focuses solely on aggregated patterns of textual and visual content at heritage sites, which aligns with established guidelines for the ethical use of user-generated content in academic research. Image data were used strictly for non-commercial, analytical purposes, and no attempt was made to identify individuals appearing in them.

3.7.2. Limitations of the Materials

The dataset reflects only the content that users choose to upload, and is therefore subject to several forms of bias. Review and image volumes vary substantially across the heritage components, with highly visited sites generating large amounts of material while less well-known locations remain partially represented. This imbalance is a structural characteristic of user-generated content and may affect the representativeness of findings. In addition, review texts contain linguistic noise, platform-specific artefacts, and variable levels of detail, while images may include irrelevant or low-information content. Although preprocessing steps were designed to address these issues, the materials still reflect the uneven and selective nature of public online engagement with heritage sites.

4. Methodology

4.1. Research Design and Analytical Overview

This chapter describes how the prepared data are turned into indicators for later analysis. The work builds on the materials and preprocessing steps described in Chapter 3, which cover data collection, sentence segmentation, language detection, and machine translation into English.

Throughout this thesis, user-generated content (UGC) is used as an umbrella term referring to publicly shared textual reviews and user-generated images collected from online platforms. The study uses two main types of user-generated content: text reviews and user photos. From these two sources, three kinds of signals are derived: experiential signals, OUV signals, and visual themes. Each signal type answers a different part of the research questions but follows the same basic data structure.

- **Experiential signals** come from the original-language review texts. Sentiment analysis provides information about how visitors evaluate their experiences at each heritage component. Keyterm extraction identifies the most frequent and characteristic words and phrases in the reviews. These two procedures outline what visitors feel and what they talk about, before any reference to UNESCO OUV criteria. Work at this stage keeps the source languages and does not rely on translated text, in order to reduce additional noise from machine translation.
- **OUV signals** come from an English sentence corpus. Chapter 3 has already introduced how reviews were split into sentences, how languages were detected, and how non-English sentences were translated into English. In the methodology, this sentence-level English corpus serves as the input to a two-stage OUV alignment pipeline. Stage-1 applies a generic UNESCO OUV classifier and produces predicted OUV criteria for each sentence. Stage-2 refines these outputs through a Macao-specific semantic alignment with a local OUV knowledge base. Together, these two stages identify which sentences in the reviews are closely related to criteria ii, iii, iv, and vi as they apply to the Historic

Centre of Macao.

- **Visual themes** come from the user photos linked to the same heritage components. Each image is first converted into a numeric representation using a pre-trained image embedding model. These image embeddings are then grouped into clusters. The clusters are interpreted as recurring visual themes, such as specific building types, streetscapes, interiors, or open spaces. These visual themes describe what visitors choose to photograph and how this focus differs between components.

All analyses follow the same hierarchy of units. The smallest textual unit is the sentence. Sentences are linked back to their reviews, and each review is linked to a heritage component. Components together form the entire World Heritage site. Image data follow the same component structure through their component identifiers. This hierarchy allows comparisons at several levels: within individual components, between different components, and at the level of the whole site.

Sentence-level and image-level outputs from the textual and visual pipelines are aggregated by `poi_id` to produce profiles for each component (for example sentiment distributions, OUV criteria proportions, and visual theme compositions). These component-level profiles can then be compared across the Historic Centre and mapped in the Results chapter to highlight spatial contrasts between different parts of the site.

4.2. Place-Based Data Organization and POI Definition

This study adopts a place-based data organization strategy grounded in a predefined set of heritage Points of Interest (POIs). The POIs correspond to the 27 officially recognized components of the Historic Centre of Macao as defined in UNESCO nomination and management documents. Each POI serves as a stable spatial anchor for data collection and subsequent analysis across platforms, languages, and modalities.

Online reviews and images are treated as place-anchored signals: each item is associated with a specific heritage component. As the analytical focus lies on public perception, narrative construction, and multimodal representation, place references are treated as fixed inputs that enable consistent aggregation and comparison.

To support this process, a lightweight local place register was compiled for the Historic Centre of Macao. For each heritage component, the register records a unique site identifier, geographic coordinates and a curated set of name variants in Chinese,

Portuguese, and English, provided in Appendix A.1. This reflects Macao’s historically layered toponymy and ensures consistent data aggregation across platforms that may use different naming conventions. The scope of this register is intentionally limited to officially designated heritage sites and does not aim to capture all businesses or minor landmarks within the historic core.

This place-based framework underpins all subsequent analytical steps. It guides the retrieval of user-generated text and images from Google Maps, TripAdvisor, and Ctrip, ensures that all data are linked to a common spatial reference, and enables systematic comparison between public narratives and officially defined OUV claims at both site-specific and aggregate levels.

4.3. Textual Analysis I: Sentiment and Keyterms

Textual analysis in this step focuses on the original languages of the reviews. The aim is to describe how visitors feel and what they talk about before any use of the UNESCO OUV framework. All methods in this section work directly on English, Simplified Chinese, and Traditional Chinese texts, in order to avoid extra noise in the emotional and lexical signals.

4.3.1. Sentiment Analysis

Sentiment analysis was applied to all visitor reviews to summarise users’ emotional evaluations across sites, languages, and platforms. Because the dataset contains English, Simplified Chinese, and Traditional Chinese comments, a multilingual model was required. A multilingual BERT sentiment classifier based on XLM-RoBERTa was adopted, since such models provide strong cross-lingual semantic representations and perform well on short, informal texts like online reviews (Devlin et al., 2019; Pires et al., 2019). The choice of a multilingual transformer is important in the context of Macao, where visitor narratives often mix English with Simplified and Traditional Chinese and sometimes include code-switching within a single review. Models such as XLM-RoBERTa provide a shared cross-lingual semantic space and have been trained on many languages simultaneously, which makes them more robust than separate monolingual models when handling mixed scripts and borrowings across languages. This cross-lingual alignment is essential for obtaining sentiment scores that are comparable across English, Simplified Chinese, and Traditional Chinese reviews.

Preprocessing for sentiment classification was kept minimal so that the tone and style of the original comments remained visible. HTML tags, emojis, and repeated whitespace were removed, while punctuation was preserved. Text encoding was normalised across Google Maps, TripAdvisor, and Ctrip, and all reviews were merged into a single multilingual dataset in which identifiers such as `poi_id`, `platform`, and `language` were retained. Sentiment analysis was carried out at the review level rather than at the sentence level, because emotional polarity in visitor feedback is usually expressed across the full text of the review.

The fine-tuned multilingual model produced three sentiment-related fields for each review: `sentiment_label`, `sentiment_score`, and `sentiment_star`.

The field `sentiment_label` stores the predicted sentiment class (positive, neutral, or negative).

The field `sentiment_score` stores a continuous score from the classifier, which reflects the strength of the underlying sentiment prediction.

The field `sentiment_star` stores the same information on an ordinal scale with discrete values from one to five.

These three variables were written back into the main review table and form the basis for descriptive summaries and later cross-analyses with OUV alignment.

4.3.2. Keyterm Extraction

Keyterm extraction was used to identify the most characteristic words and short phrases in the visitor reviews. The aim was to obtain concrete lexical units that indicate what visitors describe, value, or criticise at different sites and on different platforms.

Because almost every review mentions one or more site names, place names needed specific handling. A dedicated stopword list was created that contains all heritage components in English, Simplified Chinese, and Traditional Chinese, together with common spelling variants and forms with or without brackets or punctuation. Any candidate phrase that included one of these names was removed from the results. Removing all component names from the candidate keyphrases avoids trivial results dominated by place names but may also remove expressions where the name carries additional experiential meaning (for example, phrases such as “Take photos at the Ruins of St. Paul’s”). This potential information loss is acknowledged as a limitation of the keyterm analysis and is revisited in the Discussion.

Initial experiments with simple frequency counts and TF-IDF weighting produced unsatisfactory outputs. Lists were dominated by function words, single-character Chinese tokens, and fragmented n -grams, which is a common problem in multilingual user-generated content. To obtain more meaningful phrases, a semantic, embedding-based method was therefore adopted.

Keyterm extraction was carried out with KeyBERT, a BERT-based unsupervised method that ranks candidate phrases by their semantic similarity to the document in which they occur, which outperformed traditional approaches on short text (Khan et al., 2022; Issa et al., 2023). In the configuration used here, KeyBERT relies on the multilingual SentenceTransformer model *paraphrase-multilingual-mpnet-base-v2* (Reimers and Gurevych, 2019b). Using a multilingual SentenceTransformer model within KeyBERT likewise ensures that semantically similar expressions in different languages are treated consistently in the keyterm rankings. This procedure favours phrases that are semantically central to the review, rather than simply frequent, and works consistently across English, Simplified Chinese, and Traditional Chinese.

Keyphrases were extracted in several configurations that follow the main analytical views of the thesis. At the component level, keyterms were computed separately for each heritage component. This allows direct comparison of what visitors talk about at different sites. At the platform level, keyterms were also generated separately for Google Maps, TripAdvisor, and Ctrip. This makes it possible to see whether the same site is described differently on different platforms. At the language level, keyterms were extracted for English, Simplified Chinese, and Traditional Chinese reviews, so that language-specific patterns in visitor narratives can be examined. Ruins of St Paul's serves as an anchor case in later chapters, where these component-, platform-, and language-specific keyterms are interpreted in more detail.

For each configuration, the resulting keyphrases were stored as ranked lists and exported as CSV files. The top-ranked terms were also visualised as multilingual word clouds and as Top- N bar charts, which provide an intuitive overview of the most frequent and distinctive expressions used by visitors.

4.3.3. Cross-Platform and Cross-Language Experiential Profiles

The outputs from sentiment analysis and keyterm extraction together form a set of experiential profiles. For each platform and language group, it is possible to describe:

- how visitors feel (the share of positive, neutral, and negative reviews)

- what visitors talk about (the main keyterms and phrases).

Keyterm rankings from different platforms (Google Maps, TripAdvisor, Ctrip) and different languages (English, Simplified Chinese, Traditional Chinese) can be compared directly. In addition to visual inspection of word clouds and bar charts, pair-wise statistical comparisons were carried out on the ranked keyterm lists. For selected pairs of platforms or languages, the shared keyterms were identified and their ranking positions were compared using the non-parametric Mann–Whitney U test. This test evaluates whether the distribution of ranks for a given set of keyterms tends to be systematically higher or lower in one group than in another, without assuming normality of the underlying scores. The resulting test statistics and p -values provide a numerical summary of how similar or different the thematic focus is across platforms and languages, and these values are reported and interpreted in the Results chapter.

Taken together, these sentiment and keyterm profiles describe visitor experience in the visitors’ own words and emotional tones. They provide an experiential baseline against which the more formal OUV-related analyses in later sections can be understood.

4.4. Textual Analysis II: OUV Alignment Pipeline

This section describes how review sentences were linked to the Outstanding Universal Value (OUV) of the Historic Centre of Macao. The analysis uses a two-stage pipeline that builds on the English sentence corpus described in Section 3.5. Stage 1 applies generic UNESCO OUV classifiers from the WHOSe Heritage project (Bai et al., 2021). Stage 2 applies a Macao-specific semantic alignment between visitor sentences and a set of local OUV claims using a Sentence-BERT (SBERT) model (Reimers and Gurevych, 2019b). To ensure that the final thresholds are transparent and data-driven, both steps are checked using human interpretation.

4.4.1. Stage 1: Generic OUV Criterion Classification

Stage 1 provides a coarse filter for OUV-related content. It uses two pretrained text classification models from WHOSe Heritage (Bai et al., 2021): a BERT-based classifier and a ULMFiT-based classifier. The goal is to combine their outputs into a single high-recall decision rule that removes clear noise while retaining almost all potentially OUV-relevant sentences.

For the BERT model, the predicted OUV criterion with the highest probability is taken as the top-1 label, and the associated probability is stored as `bert_top1_prob`. For the ULMFiT model, the three most probable criteria and their probabilities are retained. The sum of these three probabilities is stored as `ulm_top3_prob` and acts as a measure of overall ULMFiT confidence in its top-3 predictions. The empirical distributions of these three signals are shown in Figure 4.1, together with the thresholds used in the final decision rule.

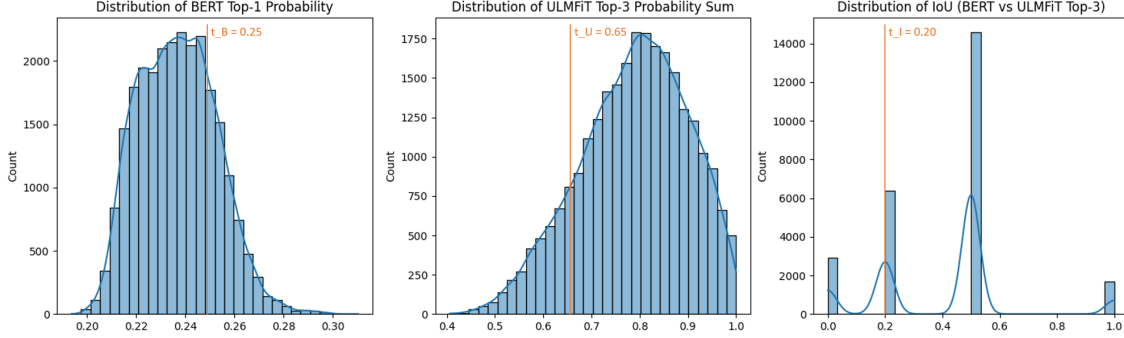


Figure 4.1.: Distributions of the three Stage 1 decision signals across all sentences: Vertical dashed lines indicate the thresholds used in the final Stage 1 rule ($t_B = 0.25$, $t_U = 0.65$, $t_I = 0.20$).

In addition to these two confidence scores, a simple agreement measure between the models is computed. For each sentence, the three most probable criteria from BERT are collected as a set, and the three most probable criteria from ULMFiT are collected as a second set. The intersection-over-union (IoU) between these two sets is then calculated and stored as `iou_top3`. Because each set can contain at most three criteria, IoU values are discrete. An IoU of zero indicates no shared criteria, whereas an IoU of at least 0.20 indicates that the two models share at least one criterion in their top-3 lists. This agreement signal is used to identify cases where ULMFiT is confident and broadly consistent with BERT (Bai et al., 2022a, 2023).

For Stage 1, 90 sentences were selected using stratified sampling based on the agreement measure `iou_top3`. Thirty sentences were drawn at random from each of three IoU bands: high agreement ($\text{iou_top3} \geq 0.8$), medium agreement ($0.3 \leq \text{iou_top3} < 0.8$), and low agreement ($\text{iou_top3} < 0.3$). Because both models contribute top-3 sets, the IoU values are discrete and in practice take only the values $\{0, 0.2, 0.5, 1.0\}$. The cut points at 0.3 and 0.8 therefore simply separate low agreement (IoU = 0 or 0.2), medium agreement (IoU = 0.5), and high agreement (IoU = 1.0). This design ensures that the validation set includes sentences where the two classifiers strongly agree, partially agree, and disagree.

Each sentence received a note in the set $\{0, 2, 3, 4, 6\}$, where 0 denotes no OUV link

and other values indicate corresponding criteria ii, iii, iv, or vi. A grid search was carried out over three thresholds: the BERT top-1 probability threshold t_B , the ULMFiT top-3 probability sum threshold t_U , and the IoU threshold t_I .

Table 4.1.: Selected threshold combinations for the Stage 1 decision rule from the grid search over BERT top-1 probability (t_B), ULMFiT top-3 probability sum (t_U), and IoU (t_I). Metrics are computed on the 90 manually annotated sentences.

t_B	t_U	t_I	Precision	Recall	F1	Keep ratio
0.25	0.65	0.20	0.800	0.938	0.863	0.833
0.25	0.675	0.20	0.808	0.922	0.861	0.811
0.27	0.65	0.00	0.789	0.938	0.857	0.844
0.27	0.675	0.20	0.798	0.922	0.856	0.822
0.25	0.675	0.00	0.772	0.953	0.853	0.878

A grid search over (t_B, t_U, t_I) was carried out on the 90 annotated sentences, and precision, recall, F1, and the proportion of sentences retained (keep ratio) were computed for each combination. Table 4.1 lists a subset of the best-performing settings. Several neighbouring combinations achieve very similar F1 values, which indicates that the chosen rule is not overly sensitive to small changes in the thresholds.

The resulting Stage 1 decision rule is:

$$\text{S1_keep} = (\text{bert_top1_prob} \geq 0.25) \vee (\text{ulm_top3_prob} \geq 0.65 \wedge \text{iou_top3} \geq 0.20). \quad (4.1)$$

This rule can be read in two parts. First, a sentence is kept if the BERT classifier is sufficiently confident in its top-1 criterion. Second, a sentence is also kept if the ULMFiT classifier assigns high probability mass to its top-3 criteria and, at the same time, the top-3 predictions of both models share at least one criterion. On the annotated sample, this rule achieves recall of at least 0.94 and precision of about 0.80, which is suitable for a coarse filtering stage.

The rule is then applied to the full sentence corpus. All sentences with $\text{S1_keep} = \text{True}$ are written to a filtered dataset. This dataset serves as the input to the Macao-specific alignment in Stage 2.

4.4.2. Stage 2: Macao-Specific OUV Claim Alignment

Stage 2 performs a fine-grained alignment between visitor sentences and Macao-specific OUV claims. It uses a Sentence-BERT (SBERT) embedding model (Reimers and Gurevych, 2019b) to represent both sides of the alignment in the same vector

space. The set of OUV claims comes from the Macao OUV knowledge base introduced in Section 3.5. Each claim is a short statement that describes one aspect of criteria ii, iii, iv, or vi as they apply to the Historic Centre of Macao.

For each sentence that passed Stage 1, the following steps are carried out:

- An SBERT embedding is computed for the sentence and for each Macao OUV claim.
- Cosine similarity is computed between the sentence embedding and the embedding of every claim.
- The claim with the highest cosine similarity is identified.
- This maximum similarity is recorded as a soft relevance score for the sentence.

This yields, for each sentence, a best-matching OUV claim and an associated similarity value $\text{sim} \in [0, 1]$.

The empirical distribution of similarity scores is summarised in Table 4.2. The median similarity is about 0.36, and the middle 50% of sentences lie between roughly 0.25 and 0.45.

Table 4.2.: Empirical quantiles of cosine similarity between visitor sentences and their best-matching Macao OUV claims (Stage 2).

Quantile	Similarity
0.05	0.120
0.10	0.166
0.25	0.254
0.50	0.362
0.75	0.454
0.90	0.527

For Stage 2 validation, a new sample of 90 sentences was selected. The sample was drawn using stratified sampling across low, middle, and high similarity bands, based on the quantiles in Table 4.2. A low-similarity band (below the first quartile, $\text{sim} < 0.254$), a middle band (between the first and third quartiles, $0.254 \leq \text{sim} \leq 0.454$), and a high band (above the third quartile, $\text{sim} > 0.454$) were defined. Thirty sentences were sampled at random from each band, so that the validation set contained clearly unrelated, borderline, and strongly related cases.

Each sentence–claim pair in this validation set was then judged by hand. Human annotators again rated the strength of the link with notes in the set $\{0, 2, 3, 4, 6\}$. Accuracy was computed within each similarity band. Sentences with similarity below the first quartile showed low accuracy (about 0.37), whereas sentences in the

middle and high similarity bands both reached accuracy values around 0.93.

To choose a decision threshold for Stage 2, a sweep over similarity values from 0.10 to 0.70 was carried out. For each candidate threshold T , precision, recall, F1, and accuracy were computed with respect to the human labels. Figure 4.2 shows F1 and accuracy as functions of T . Both curves peak at approximately $T = 0.259$, and performance remains stable in a small neighbourhood of this value. The value was therefore chosen as the working threshold for Stage 2.

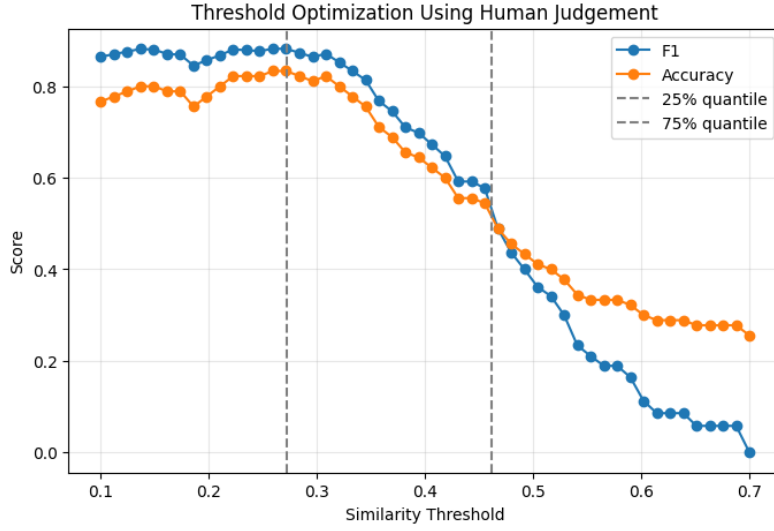


Figure 4.2.: F1 and accuracy of the Stage 2 decision rule as functions of the similarity threshold T , evaluated on 90 manually annotated sentences. Both curves peak at approximately $T = 0.259$, which is used as the working threshold.

The final Stage 2 decision rule is:

$$\text{S2_keep} = \begin{cases} \text{True}, & \text{if similarity} \geq 0.25918, \\ \text{False}, & \text{otherwise.} \end{cases} \quad (4.2)$$

Only sentences with $\text{S2_keep} = \text{True}$ are treated as aligned to the Macao OUV claims and carried forward to the final OUV-bearing dataset.

4.4.3. Final Two-Stage Filtering

The two stages are combined through a simple logical rule. Stage 1 removes sentences that are clearly unrelated to UNESCO criteria, while still keeping almost all potentially relevant material. Stage 2 then applies a stricter, Macao-specific filter based on similarity to local OUV claims. In combination, the two stages produce a dataset with a high proportion of sentences that genuinely relate to the OUV

description of the Historic Centre. Applied to the full corpus, Stage 1 retained 21145 sentences (82.7% of all sentences), and Stage 2 then marked 15646 of these as high-confidence alignments (73.99% of Stage 1 sentences).

Sentences kept were written to a final dataset. In this set, sentences from the middle and high similarity bands dominate, and human validation shows accuracy values around 0.93 for these bands. This final dataset is used in later chapters to construct OUV profiles for each heritage component, to compare Stage 1 and Stage 2 outputs, and to select representative qualitative examples. The two-stage pipeline thus provides a reproducible and human-validated way to detect OUV-related content in user-generated reviews, and it extends the WHOSe Heritage framework with claim-level alignment and a Macao-specific interpretation layer.

4.4.4. Aggregation and Visualisation

The OUV-bearing sentences from the two-stage pipeline form the basis for all subsequent OUV profiles. Each sentence in the final dataset carries three key attributes: a heritage component identifier (`poi_id`), a Stage 1 predicted criterion, and a Stage 2 best-matching criterion derived from the Macao OUV claims. These attributes allow aggregation at different spatial and analytical levels.

At the global level, sentences are grouped by OUV criterion and simple frequency counts are computed. Counts are then converted into proportions by dividing by the total number of OUV-bearing sentences. This produces a site-wide OUV profile that shows the relative contribution of criteria ii, iii, iv, and vi in visitor narratives.

At the component level, the same procedure is repeated for each heritage component. For each `poi_id`, the numbers of sentences linked to criteria ii, iii, iv, and vi are counted and normalised to proportions. This yields one OUV profile per component, which can be compared across sites. In some cases, the aggregation is further stratified by platform and language.

The relationship between Stage 1 and Stage 2 is summarised through a criterion transition table. For all sentences that pass Stage 1, a contingency table is constructed with Stage 1 criteria as rows and Stage 2 criteria as columns. Each cell records the number of sentences that move from a given Stage 1 criterion to a given Stage 2 claim criterion. Row-wise normalisation produces conditional proportions that show how Stage 2 refines or reassigns the Stage 1 predictions. These aggregated profiles are visualised using simple graphics, which will be presented and interpreted in the Results chapter.

Finally, component-level textual and visual outputs were combined to enable cross-modal comparison. For each heritage component, aggregated visual theme proportions were paired with OUV-aligned textual profiles derived from the Stage 2 classification. Spearman rank correlations were used to assess structural alignment between visual and textual signals across components, focusing on relative patterns rather than absolute values or causal relationships. This joint analysis forms the basis for the system-wide text–image comparison reported in Section 5.5.

4.5. Visual Analysis: Image Embeddings and Clustering

This section describes how user-generated photographs are transformed into a small set of visual themes. The analysis starts from a cleaned image corpus with metadata, then extracts deep visual features, applies clustering in a reduced feature space, and finally interprets and maps the resulting themes at the level of heritage components.

4.5.1. Deep Visual Feature Extraction

The visual feature extraction and embedding pipeline closely follows the approach described by Bai et al. (2023). In this setup, all photographs are processed with a VGG-16 network (Simonyan and Zisserman, 2015) pre-trained on the ImageNet dataset (Deng et al., 2009) using the Keras library. The activations of the final 4096-dimensional hidden layer are taken as structured visual representations for each image. These 4096-dimensional features are then reduced to 300 dimensions with Principal Component Analysis (PCA), and the resulting 300D vectors are used both for K-means clustering and for two-dimensional visualisation. For the latter, the 300D features are embedded into a two-dimensional space with t-distributed Stochastic Neighbour Embedding (t-SNE) as implemented in scikit-learn (van der Maaten et al., 2008), and the t-SNE coordinates are mapped onto a regular image grid using the RasterFairy Python library, again following Bai et al. (2023).

4.5.2. Global Visual Clustering with K-means

The aim of the clustering step is to obtain a small set of global visual themes that summarise recurring photographic patterns across all heritage components.

Clustering is carried out in the 300-dimensional VGG-PCA feature space using K-means.

To choose a suitable number of clusters K , a subset of 5 000 images is drawn at random from the full corpus. K-means is then run on this subset for candidate values $K \in \{6, 8, 12, 16, 20, 24, 28\}$, with multiple random initialisations for each value. This range covers coarse to moderately fine partitions of the image corpus. Very small values such as $K = 4$ would merge clearly distinct visual motifs (such as temples, churches, streetscapes, and green spaces) into a few broad clusters and are therefore not useful for interpretation. For every K , two diagnostics are recorded: the within-cluster sum of squared distances (inertia) and the average silhouette coefficient, as summarised in Figure 4.3. Inertia decreases steadily as K increases, while silhouette scores remain low but reach their highest value at $K = 8$. Because the improvements in these metrics beyond eight clusters are small and a limited number of themes is easier to interpret, $K = 8$ is selected as a practical compromise.

With $K = 8$ fixed, K-means is then fitted on all 14 079 images in the 300-dimensional feature space, using the scikit-learn implementation with multiple initialisations and a fixed random seed. Each image receives a cluster label `cluster_id` $\in \{0, \dots, 7\}$, which is added to the image metadata table. Cluster sizes range between 213 and 2 885 images (about 1.5% to 20.5% of the dataset). This range indicates that no single cluster dominates the corpus and that smaller clusters likely represent more specific, but still coherent, visual motifs.

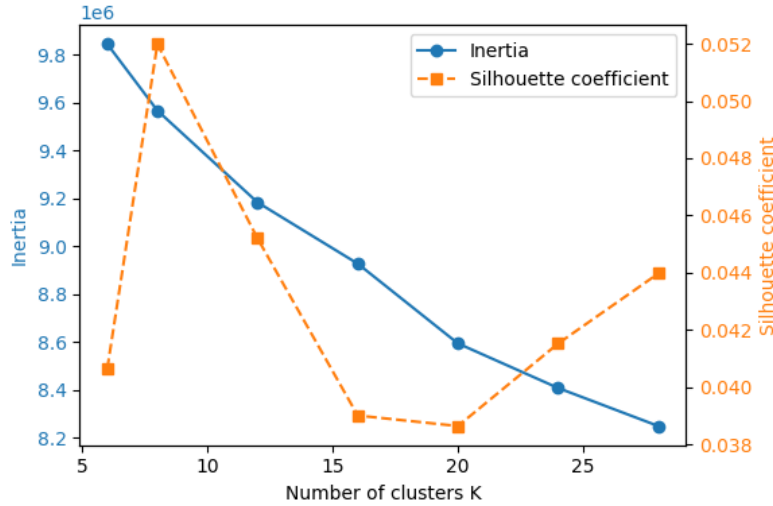


Figure 4.3.: K selection for visual clustering in the 300-dimensional feature space. The within-cluster sum of squared distances (inertia) decreases monotonically with larger K , while the average silhouette coefficient peaks at $K = 8$. Given the limited improvements beyond eight clusters and the need for interpretable themes, $K = 8$ is chosen as the final number of visual clusters.

4.5.3. Qualitative Interpretation and Naming of Visual Themes

K-means produces numerical clusters, but their semantic meaning must be interpreted. To do this, contact sheets are generated for each cluster. For each cluster c , up to 100 images with `cluster_id = c` are sampled at random and read from disk. Images are resized to small thumbnails and arranged in a grid, creating a mosaic that gives a quick visual overview of the typical content of that cluster.

These mosaics are inspected to identify recurring scenes, motifs, colours, and compositions. Based on this inspection, concise descriptive labels are assigned to each of the eight clusters. For the final solution, the themes are:

- C0 – Temples & Folk Shrines,
- C1 – Buildings & Streetscapes,
- C2 – Lanes, Alleys & Walls,
- C3 – Interiors & Decorative Details,
- C4 – Guia Fortress & Lighthouse,
- C5 – Green Spaces & Parks,
- C6 – Catholic Church Interiors,
- C7 – Facades.

These labels are stored in a new column (`visual_theme`) in the metadata table, so that every image is associated with both a numeric cluster identifier and a human-readable theme name. This step translates unsupervised clusters into interpretable categories that can later be compared to textual OUV dimensions.

4.5.4. Visualisation of the Global Feature Space

To show the structure of the visual feature space and how the eight themes relate to each other, a two-dimensional projection is created for a subset of images. A sample of 600 images is drawn from the corpus, and their 300-dimensional feature vectors and cluster labels are used as input.

The 300-dimensional vectors are projected into two dimensions using t-SNE (van der Maaten et al., 2008) with perplexity set to 30, PCA initialisation, 500 iterations, and a fixed random seed. t-SNE helps to preserve local neighbourhoods. Images that are close in the 300-dimensional space tend to remain close in the two-dimensional

map, even though the cluster boundaries are not perfectly separated.

The resulting t-SNE coordinates are then normalised to the unit square and passed to the RasterFairy library. RasterFairy assigns each point to a cell in a regular grid (for example 25×24), which converts the irregular scatter into a layout where each grid cell holds exactly one image. For each sampled image, a small thumbnail is created, converted to grayscale, and blended with a solid colour that indicates its cluster label. These tinted thumbnails are placed at the grid positions given by RasterFairy, and a legend shows the eight colours together with their theme labels.

The resulting mosaic is used as an exploratory visualisation of the image corpus. The cluster assignments themselves come from K-means in the 300-dimensional feature space; t-SNE is used only to provide an interpretable two-dimensional display of how the visual themes overlap and transition.

4.5.5. Visual Theme Distributions per Heritage Component

Finally, the distribution of visual themes is summarised for the whole site. Using the full labelled metadata table, the number of images per component and theme is counted. For selected sites, such as the Ruins of Saint Paul's, the theme proportions are plotted as bar charts to show each component's visual fingerprint across the eight themes. These fingerprints are later compared with the OUV-related textual profiles derived from user reviews. In this way, the image analysis pipeline turns unstructured user photographs into a small set of interpretable visual themes, a map of how these themes are arranged in feature space, and per-site theme distributions that can be linked back to UNESCO OUV criteria and to visitor narratives.

4.6. Anchor Case Design: Ruins of St Paul's

Ruins of St Paul's is treated as an anchor case in this study. Within the Historic Centre, it has an iconic stone façade, attracts the largest number of reviews and images in the dataset, and occupies a central position in typical visitor itineraries through the Historic Centre, which makes it a natural anchor for interpreting system-wide spatial patterns. Figure 4.4 shows its location within the Historic Centre, and Figure 4.5 illustrates the façade that appears both in the official documentation and in many user-generated images. Methodologically, Ruins of St Paul's provides a rich and recognisable example that is used to develop, tune, and illustrate the analytical pipeline before applying it to all heritage components

Compared to the other heritage components, Ruins of St Paul’s is analysed at a finer level of detail. For this component, sentiment analysis is broken down more explicitly by platform and language, keyterm rankings are compared across platforms and languages, and Stage 2 aligned sentences are examined more closely. On the image side, the component’s distribution across the eight visual themes is inspected in more depth and related to its textual OUV profiles, for example the dominance of the “Facades” cluster.

In the Results and Discussion chapters, Ruins of St Paul’s offers a concrete case where experiential themes in visitor narratives (such as crowding, photography, nearby snacks, and atmosphere) and OUV-related themes (such as architecture, East–West fusion, and religious heritage) can be directly compared. This anchor case helps to link the quantitative indicators back to a familiar and highly visible site within the World Heritage property.

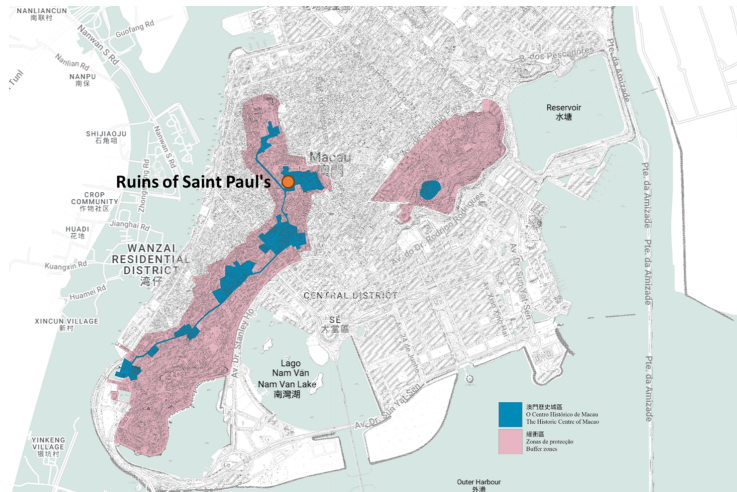


Figure 4.4.: Location of Ruins of St Paul’s within the Historic Centre of Macao. Cited from the official map (Cultural Institute of the Macao SAR Government, 2005).



Figure 4.5.: Ruins of St Paul's facade in the Historic Centre of Macao. Photograph reproduced from the UNESCO World Heritage Centre (Cultural Institute of the Macao SAR Government, 2005).

4.7. Cross-modal synthesis strategy

Cross-modal synthesis is implemented through a shared unit of analysis (Fig. 4.6): the officially designated heritage component (linked via `poi_id`). Outputs from the three analytical streams are aggregated to this level so they can be compared directly. Textual analysis yields component-level sentiment and keyterm profiles (computed in the original languages and stratified by platform and language). OUV alignment yields Macao-specific criterion profiles from the two-stage pipeline applied to translated English sentences. Image analysis yields component-level compositions of global visual themes derived from clustering.

Synthesis is carried out by comparing these component profiles across dimensions and by examining how patterns shift across platforms and language groups. Alignment between dimensions is assessed using rank-based association measures (e.g., Spearman correlations) and interpreted in terms of convergence and divergence. The synthesis outputs are reported as component profiles and comparative visual summaries that link experiential narratives, OUV framings, and dominant visual motifs.

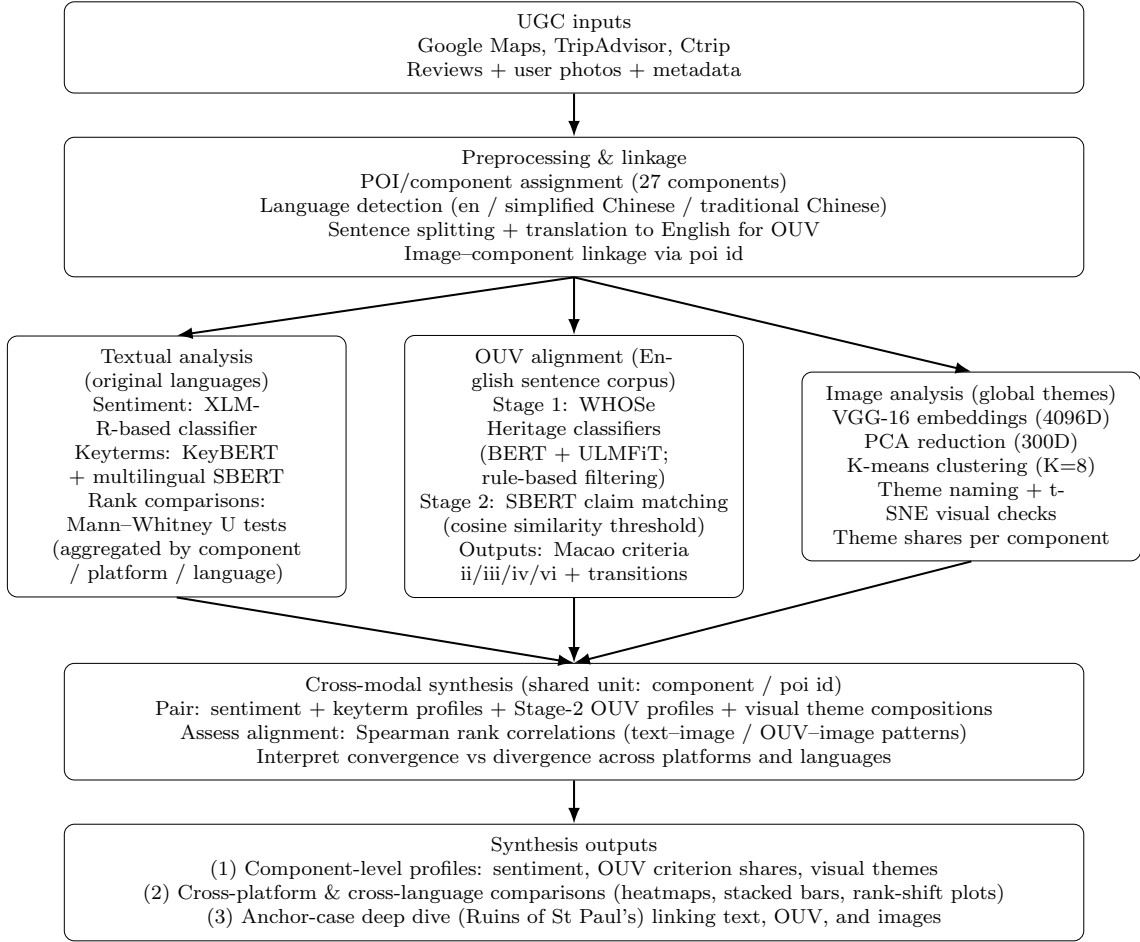


Figure 4.6.: Conceptual synthesis across the three analytical dimensions.

5. Results

5.1. Overview of Outcomes

This chapter reports what the analysis pipeline produced, and what these outputs reveal about how user-generated content reflects the ways visitors describe and visually represent the Historic Centre of Macao across platforms and languages. The results are organized in three steps: (1) textual patterns, (2) visual patterns, and (3) their joint interpretation. To make the workflow concrete before moving to system-wide summaries, the chapter first presents an anchor case (Ruins of St. Paul’s), and then generalizes to all components.

Across the full dataset, the main analytical outcomes are a set of derived, standardized outputs that can be compared across components, platforms, and languages. Concretely, the pipeline produced:

- **A sentence-level textual dataset aligned to OUV claims.** Reviews were segmented into sentences and filtered in two stages. Stage 1 retained 21,145 sentences (with an additional 4,422 sentences excluded at this stage; 82.7% retained). Stage 2 then selected 15,646 sentences as high-confidence aligned, corresponding to 73.99% of the Stage 1 retained set.
- **A global image theme assignment.** A total of 14,079 images were assigned to one of eight visual clusters (themes), enabling component-level comparisons of dominant visual representations.
- **Component profiles for cross-modal comparison.** For each heritage component, the analysis aggregates textual signals (sentiment distributions and OUV-aligned sentence shares) and visual signals (theme distributions). These profiles allow side-by-side comparison of components with very different data volumes.

Two points are important for interpreting the figures in the remainder of this chapter. First, the results are comparative rather than exhaustive. The goal is not to claim what a site is, but to describe systematic differences in how it is talked about and

depicted across platforms and languages under a consistent processing and classification pipeline. Second, coverage is uneven across components, so component-level plots should be read alongside the corresponding data volumes. Components with limited text or images are treated as lower-evidence cases.

The rest of the chapter follows this logic. Section 5.2 illustrates the full range of analytical outputs on the anchor case (Ruins of St. Paul’s), including data coverage, sentiment and keyterms, OUV-aligned sentences, and visual themes. Sections 5.3 and 5.4 then summarize system-wide textual and visual patterns, respectively. Finally, Section 5.5 brings both modalities together by comparing component-level profiles to examine structural alignments between visual themes and OUV-related narratives, and closes with a compact summary of the main empirical findings.

5.2. Anchor Case: Ruins of St Paul’s

This section uses Ruins of St Paul’s as an anchor case to show the full analytical workflow on a single, high-visibility component. Presenting one component in depth helps clarify how the textual and visual outputs relate to each other, before moving to system-wide comparisons. All results reported here follow the same preprocessing, sentiment analysis, keyterm extraction, OUV-alignment, and image-theme assignment procedures described in the Methodology chapter.

5.2.1. Data Coverage

Ruins of St Paul’s has the highest data volume in this study in both modalities. It is consistently well represented across the review platforms and also has the largest associated image set, making it an informative case for illustrating the distributional patterns that later appear in the system-wide results.

Data coverage is summarized along two dimensions: (1) the number of reviews available for this component by platform and language, and (2) the number of images linked to this component and included in the visual clustering analysis. These counts are reported in Appendix A.2.

5.2.2. Sentiment and Keyterms

Figure 5.1 summarizes mean sentiment scores for Ruins of St Paul’s by platform and language on a 1–5 scale. Across all observed platform–language combinations,

sentiment is generally positive (all means are above 3.7). English reviews show the highest averages in this anchor case, with TripAdvisor (English) at 4.24 and Google Maps (English) at 4.12. Chinese reviews are slightly lower on average, with Google Maps at 3.88 (Simplified Chinese) and 3.81 (Traditional Chinese), and Ctrip at 3.71 (Simplified Chinese). Empty cells indicate that this specific platform–language combination was not present in the dataset for this component.

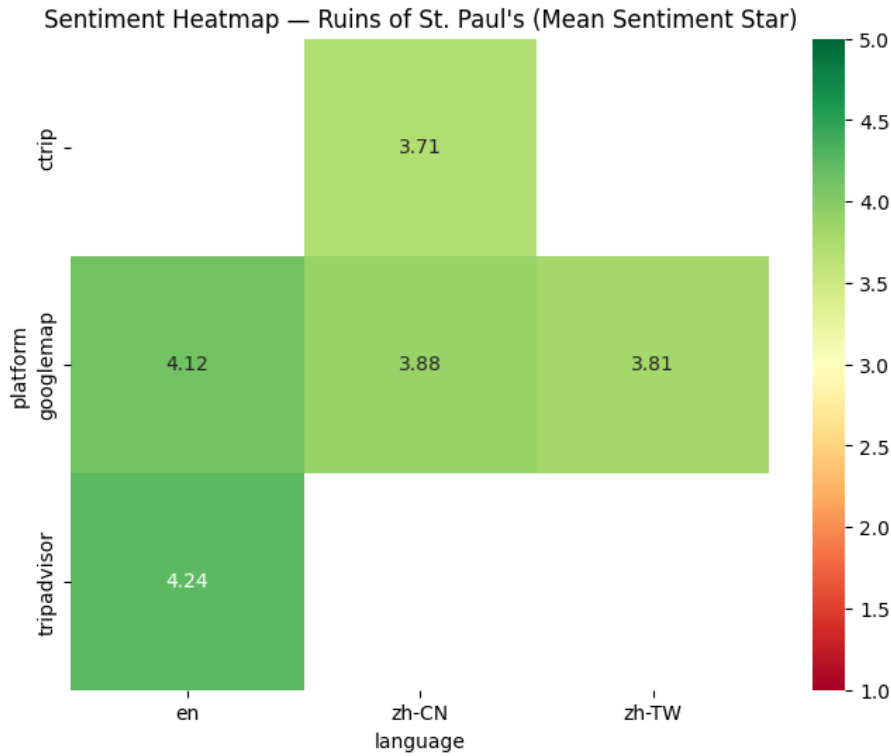
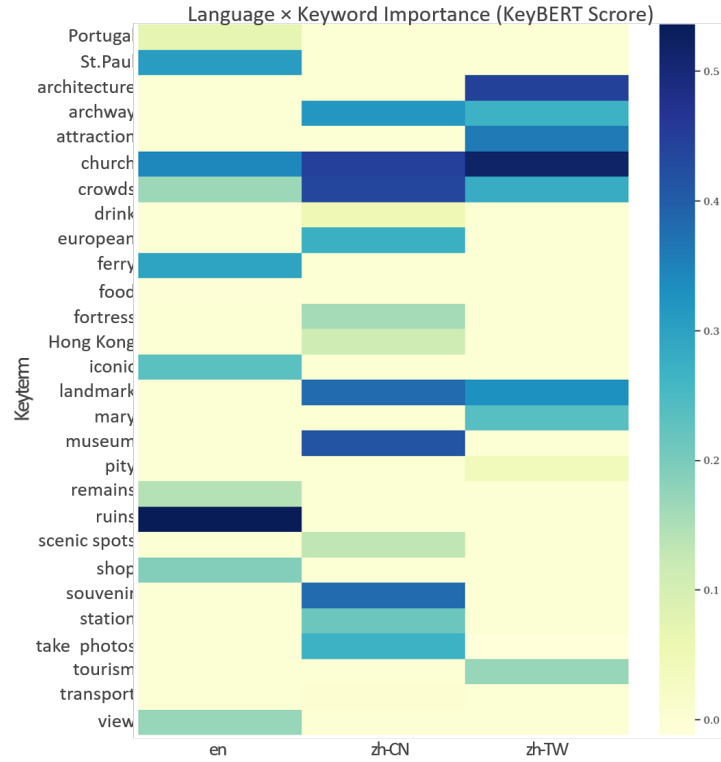
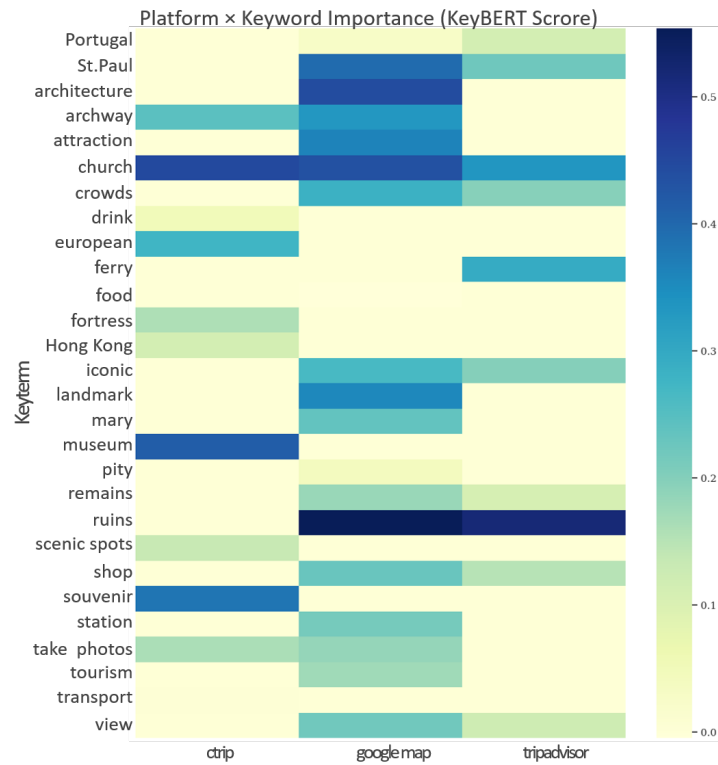


Figure 5.1.: Ruins of St Paul’s: mean sentiment (1–5) by platform and language. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.

Keyterms provide a complementary view of what visitors emphasize when describing the site. The multilingual word cloud in Figure 5.2 is dominated by terms referring to (1) the landmark identity and built form of the site (e.g., *ruins*, *church*, *landmark*, and the Chinese term 牌坊 /*archway*), and (2) the visit experience in a highly touristic setting (e.g., *crowds*, 拍照 /*take photos*, and *shop*). Additional terms point to cultural framing and surrounding context (e.g., *Portugal*, *European*, *tourism*, and practical references such as *station/transport*).



(a) Language × keyword importance. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.



(b) Platform × keyword importance.

Figure 5.3.: Ruins of St Paul's: keyterm importance heatmaps comparing (a) languages and (b) platforms. Higher values indicate stronger keyword importance within the corresponding subset.

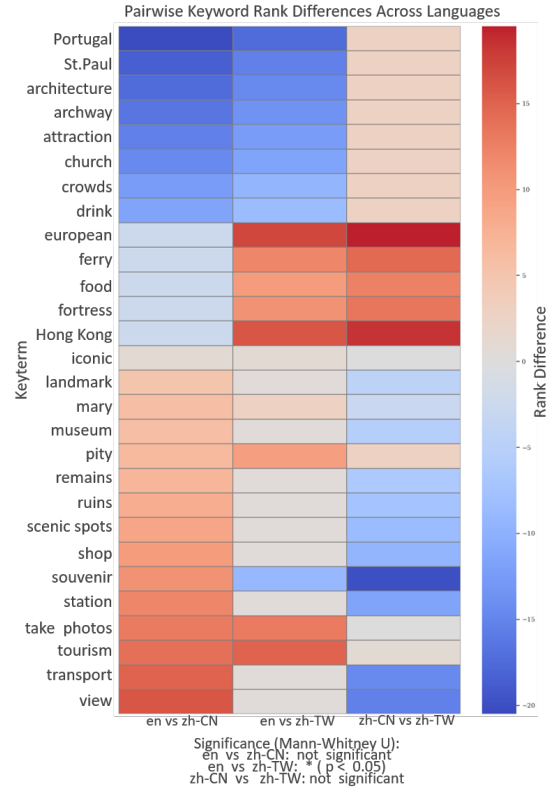
as *ruins* and *St.Paul*, and visit-experience terms such as *crowds* and *take photos*). TripAdvisor emphasizes the landmark identity and the viewing experience (notably *ruins*, *church*, and *view*), while Ctrip shows stronger prominence of visit-context terms (e.g., *museum* and *souvenir*) together with core descriptors. These differences foreshadow the system-wide platform patterns reported later in Section 5.3.

To move beyond visual inspection of KeyBERT scores, pairwise statistical tests were used to evaluate keyword rankings across subsets. Figures 5.4b report, for each keyword, the difference in its position between two ranked lists (shown as Rank for each pairwise comparison). Negative values indicate that a keyword appears earlier (i.e., is higher-ranked) in the first group of the pair, while positive values indicate a higher rank in the second group.

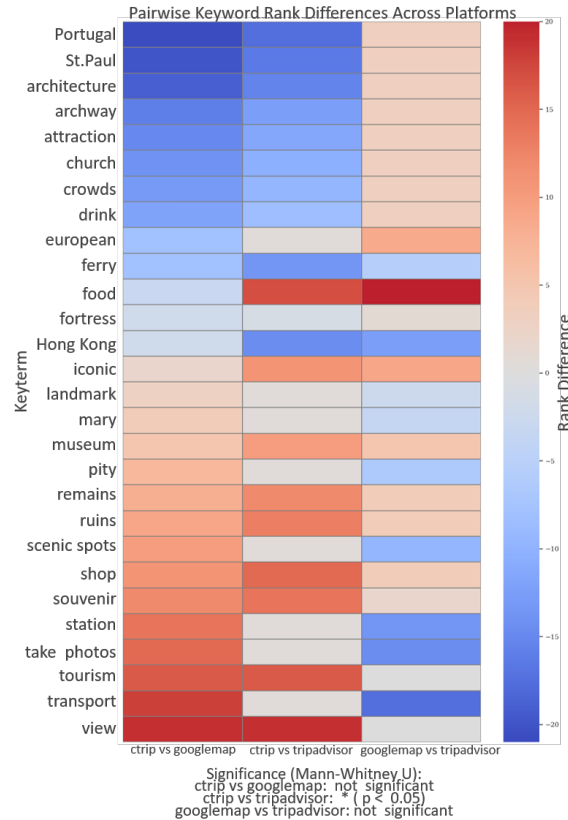
Across languages (Figure 5.4a), the overall rank distributions differ significantly only for English versus Traditional Chinese (Mann–Whitney U, $p < 0.05$). At the keyword level, site-identity terms such as *ruins* and *St.Paul* tend to be higher-ranked in the English subset, while built-form and interpretive descriptors (e.g., *architecture*, *attraction*) show stronger prominence in the Traditional Chinese subset. The keyword *archway* shows a clear rank shift toward the Chinese subsets, consistent with the linguistic note that 牌坊 frequently appears as part of the site name 大三巴牌坊, and is further reinforced by the arch-like visual form of the facade.

Across platforms (Figure 5.4b), only Ctrip versus TripAdvisor shows a significant difference in rank distributions (Mann–Whitney U, $p < 0.05$). The pattern suggests that Ctrip places relatively more emphasis on visit-context terms (e.g., *museum*, *souvenir*), whereas Google Maps and TripAdvisor rank site-identity and description terms (e.g., *ruins*, *St.Paul*, *architecture*) more prominently.

Overall, the pairwise tests support the descriptive heatmaps by indicating that the strongest systematic differences for this anchor case occur (1) between English and Traditional Chinese, and (2) between Ctrip and TripAdvisor.



(a) Pairwise keyword rank differences across languages. Language codes follow the dataset convention: en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.



(b) Pairwise keyword rank differences across platforms.

Figure 5.4.: Ruins of St Paul's: pairwise differences in keyword rank positions.

5.2.3. OUV-Aligned Sentences

This subsection reports the OUV-alignment outcomes for Ruins of St Paul’s, focusing on the Macao-specific OUV alignment outputs, and then contrasts them with the upstream generic OUV classification. All threshold selection and validation steps are described in Chapter 4.4. Here the focus is on the empirical patterns produced for the anchor case.

Figure 5.5 shows the overall Macao-specific OUV criterion profile for this component (reported as shares). The aligned discourse is strongly dominated by Criterion (iv) (*material urban fabric & architectural ensembles*), with Criterion (iii) (*contact zone & cultural fusion*) as a secondary contribution. Criterion (ii) (*cultural–intellectual exchange mechanisms*) and Criterion (vi) (*circulation of ideas & socio-political change*) appear only as small residual shares in this anchor case. In other words, once restricted to high-confidence Macao-specific alignments, visitor descriptions of Ruins of St Paul’s are most often expressed in terms consistent with the (iv)/(iii) framing rather than (ii)/(vi).

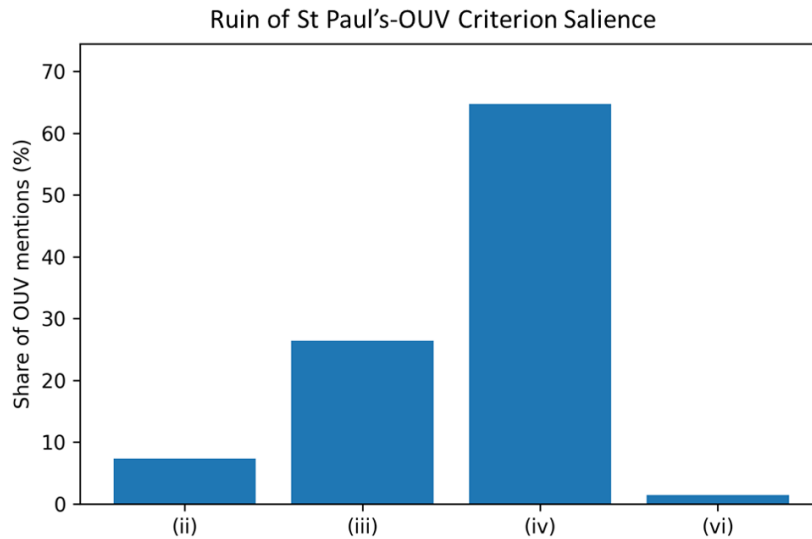


Figure 5.5.: Ruins of St Paul’s: Macao-specific OUV alignment criterion profile.

To examine whether this profile depends on language or platform, Figure 5.6 breaks the same Macao-specific aligned set down by language and by platform. The criterion composition is broadly stable across subsets: Criterion (iv) (*material urban fabric & architectural ensembles*) remains the largest share in all three language groups and across the three platforms. Differences that do appear are modest in magnitude and mainly reflect re-weighting between Criterion (ii) (*cultural–intellectual exchange mechanisms*) and Criterion (iii) (*contact zone & cultural fusion*), while Criterion (vi) (*circulation of ideas & socio-political change*) stays consistently small.

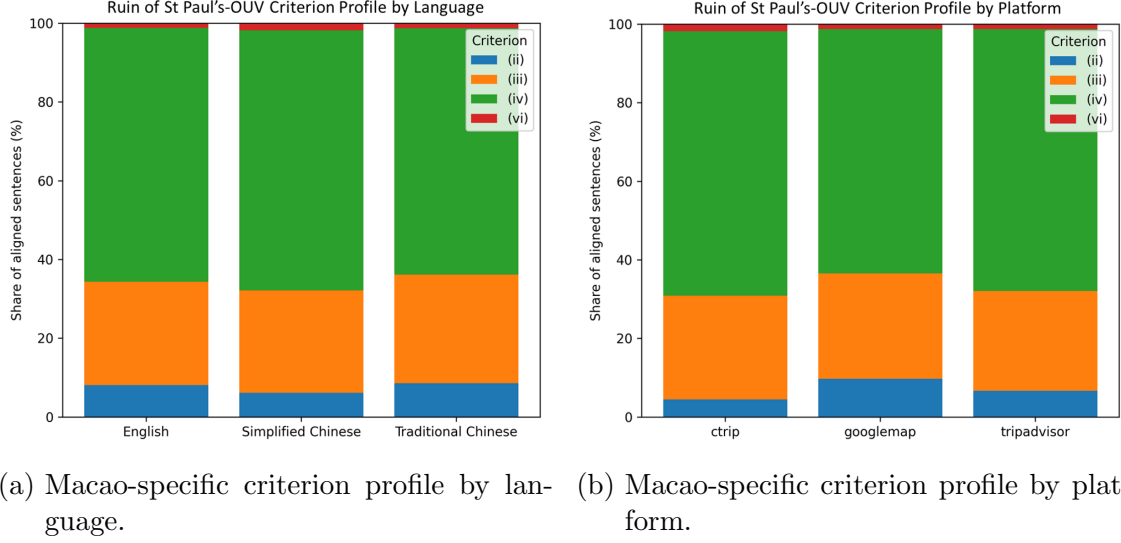


Figure 5.6.: Ruins of St Paul's: Macao-specific OUV alignment criterion composition across (a) language groups and (b) platforms.

A clearer contrast emerges when comparing the two stages directly. Figure 5.7(a) shows that the generic OUV classification output for this component is heavily concentrated on Criterion (ii) (*cultural-intellectual exchange mechanisms*), whereas the Macao-specific OUV alignment output shifts strongly toward Criterion (iv) (*material urban fabric & architectural ensembles*) and Criterion (iii) (*contact zone & cultural fusion*). The sentence-level transition matrix in Figure 5.7(b) makes this re-mapping explicit: conditional on each generic label, the majority of sentences are assigned to criterion (iv) in the Macao-specific alignment (e.g., 64.1% for generic(ii), 80.0% for generic(iii), 67.7% for generic(iv), and 100.0% for generic(vi)). Smaller portions transition into Macao-specific (iii), and only minor shares remain in (ii) or (vi). Taken together, these plots indicate that the Macao-specific alignment not only reduces volume but also substantially re-weights the criterion composition relative to the generic classification.

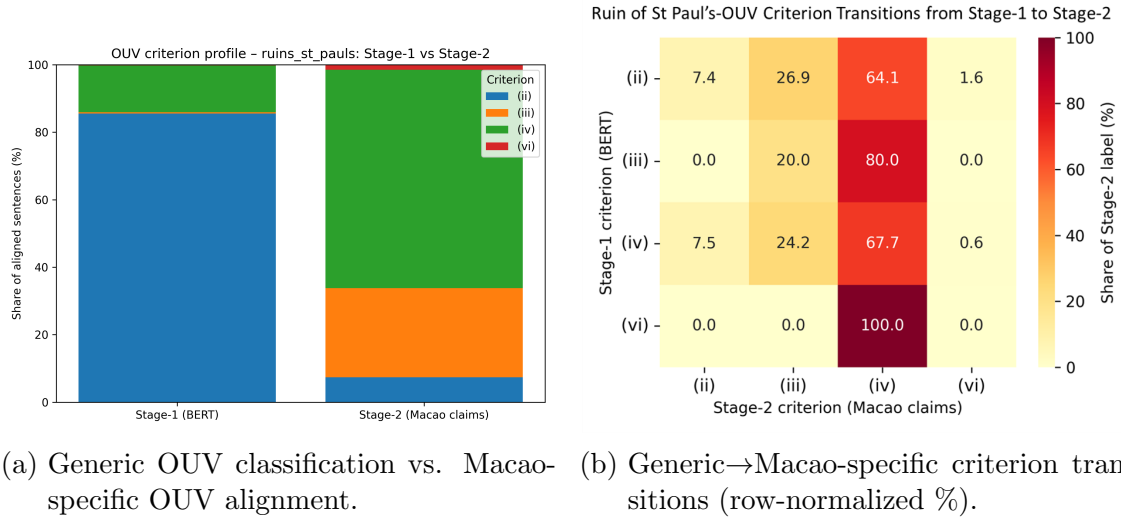


Figure 5.7.: Ruins of St Paul's: how Macao-specific OUV alignment refines generic OUV classification. Panel (b) is row-normalized: within each generic label (row), values sum to 100% across Macao-specific labels (columns).

5.2.4. Visual Themes

Visual themes were derived from an unsupervised clustering analysis applied to the pooled image corpus across all heritage components. For the anchor case, the outputs therefore summarise how images of Ruins of St Paul's are distributed across the global clusters, and how these clusters can be interpreted through sampled images.

Figure 5.8 reports the theme composition for Ruins of St Paul's. The component is visually dominated by *Buildings & Streetscapes* (39.63%) and *Facades* (28.32%), followed by *Lanes, Alleys & Walls* (13.06%). Smaller shares are assigned to *Interiors & Decorative Details* (7.68%) and *Green Spaces & Parks* (6.59%), while themes that primarily represent other landmark types in the Historic Centre (e.g., *Temples & Folk Shrines*, *Catholic Church Interiors*, and *Guia Fortress & Lighthouse*) contribute only marginally for this component.

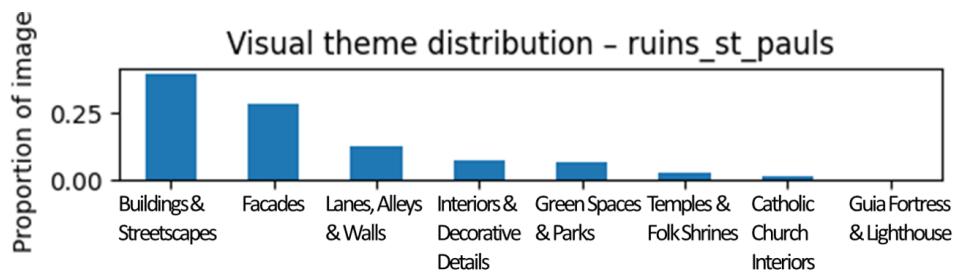
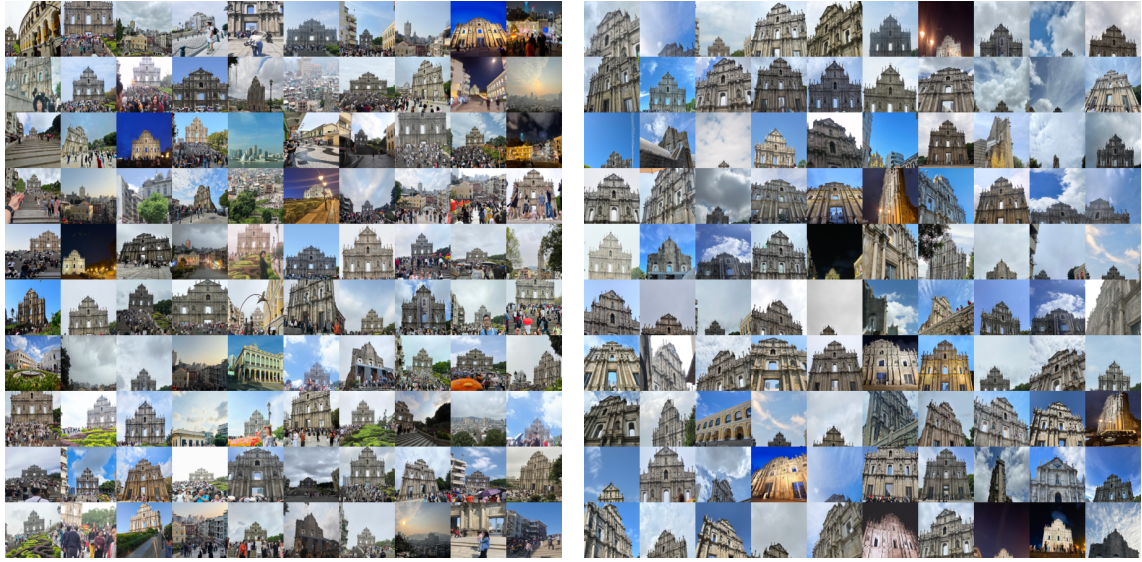


Figure 5.8.: Ruins of St Paul's: distribution of images across the eight global visual theme clusters (shares within the component).

Because Ruins of St Paul’s is a highly photographed icon with a visually distinctive facade, the two largest clusters contain many images of the same object. The sampled grids (Figure 5.9) show, however, that the clusters differ systematically in photographic framing and scene context rather than in object identity. *Buildings & Streetscapes* is dominated by wider views in which the façade is embedded in the surrounding urban scene—most commonly with the forecourt/plaza visible and often including crowds, adjacent buildings, or the street network. In contrast, *Facades* concentrates on closer, more front-on compositions that isolate the architectural surface as the main subject, with less of the plaza and surrounding context in frame. For this anchor case, the clustering therefore captures a meaningful distinction between landmark-in-context versus architectural close-up representations.



(a) *Buildings & Streetscapes*: landmark-in-context views (wider framing, surrounding streetscape/crowds).

(b) *Facades*: close-up or front-on views focusing on the façade as the primary subject.

Figure 5.9.: Ruins of St Paul’s: sampled images from the two dominant visual theme clusters.

Figure 5.10 shows the distribution of visual themes for the anchor case, the Ruins of St. Paul’s, disaggregated by image language. Across English, Simplified Chinese, and Traditional Chinese images, *Buildings & Streetscapes* and *Facades* consistently dominate, indicating a language-independent emphasis on the site’s exterior architectural presence and its immediate urban context. At the same time, modest language-specific differences are observable: English-language images place slightly greater emphasis on streetscapes, Simplified Chinese images allocate a higher share to lanes and surrounding urban elements, and Traditional Chinese images show a relatively stronger focus on interior and decorative details. Minor themes remain secondary across all languages, while the near absence of the *Guia Fortress & Light-*

house theme reflects the site-specific character of the Ruins of St. Paul’s. Overall, the anchor case suggests a stable core visual narrative accompanied by subtle language-sensitive framing preferences.

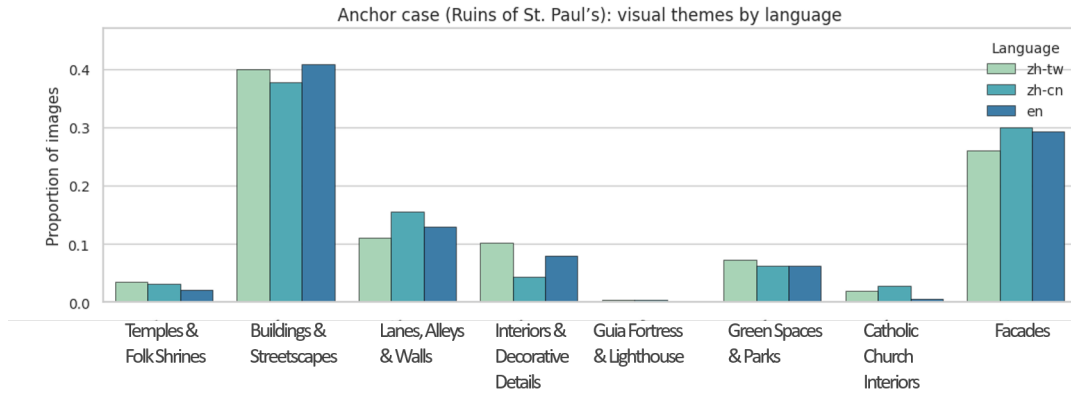


Figure 5.10.: Anchor case (Ruins of St. Paul’s): distribution of visual themes by image language. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.

5.3. System-wide Textual Patterns

5.3.1. Sentiment Across Components, Platforms, and Languages

Figure 5.11 maps the mean sentiment score (1–5) for all heritage components, showing that the overall tone of reviews is positive across the system. The map provides a compact overview of where higher and lower mean scores occur within the Historic Centre, and it also makes the unequal data presence across components visually apparent through the varying marker sizes encoded in the figure.

A cross-source comparison is provided in Figure 5.12. Across the available platform–language combinations, mean sentiment remains consistently high, with values mostly around four stars. English subsets show the highest averages on Google Maps (4.24) and TripAdvisor (4.19). Chinese subsets on Google Maps are slightly lower (3.94 for Simplified Chinese; 3.91 for Traditional Chinese). For Ctrip, the available subset (Simplified Chinese) has a mean of 4.06. Blank cells indicate platform–language combinations not present in the processed dataset.

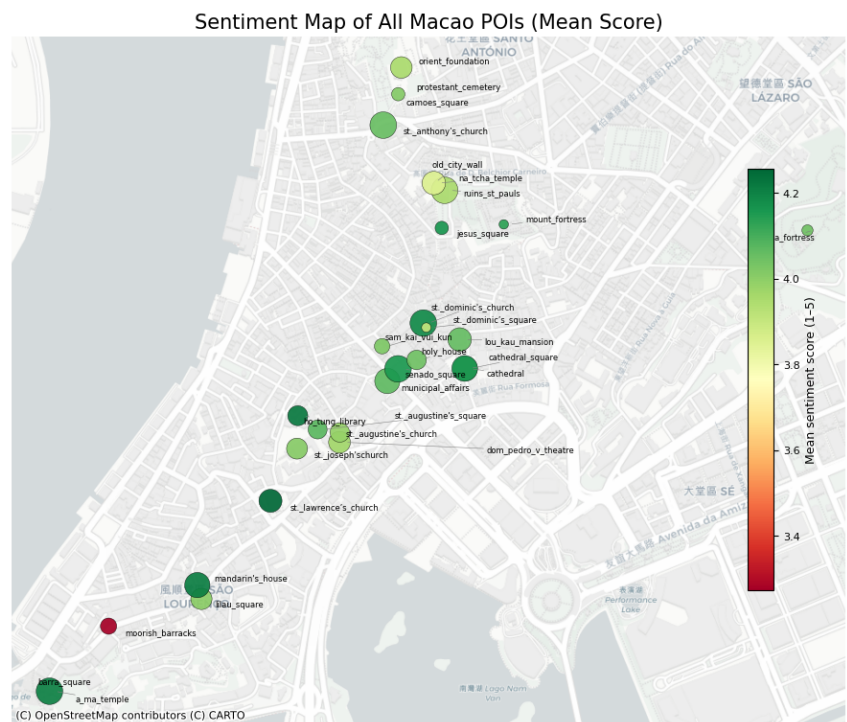


Figure 5.11.: System-wide sentiment map for the Historic Centre of Macao. Colors represent mean sentiment scores; Marker size is proportional to the number of reviews. Basemap: (C) OpenStreetMap contributors (C) CARTO.

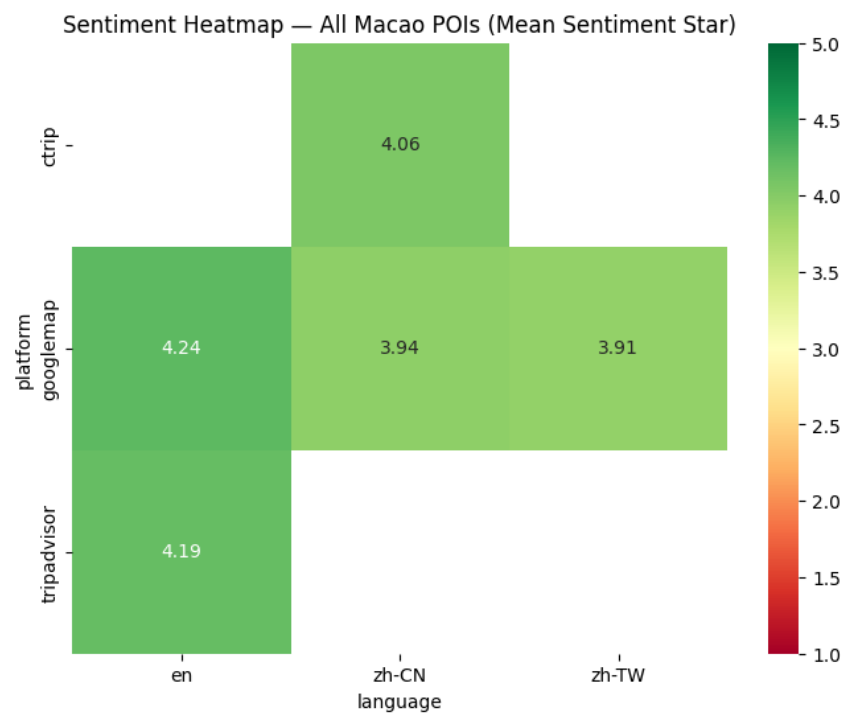


Figure 5.12.: System-wide sentiment by platform and language. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.

Component-level variation is summarised in Figure 5.13. Mean scores fall within a relatively narrow range, from 3.27 (Moorish Barracks) to 4.26 (St. Lawrence’s Church). Several components cluster at or above approximately 4.2 (e.g., Ho Tung Library 4.22; Mandarin’s House 4.21; A-Ma Temple 4.20), while a small number are notably lower (e.g., Na Tcha Temple 3.85; St. Dominic’s Square 3.90). Ruins of St Paul’s (3.95) sits close to the system-wide middle rather than at the top of the distribution, despite being the most visible anchor case in terms of data volume. Overall, the global results indicate that sentiment is broadly positive across the heritage system, with modest but interpretable differences by platform/language and a limited set of lower-scoring components.

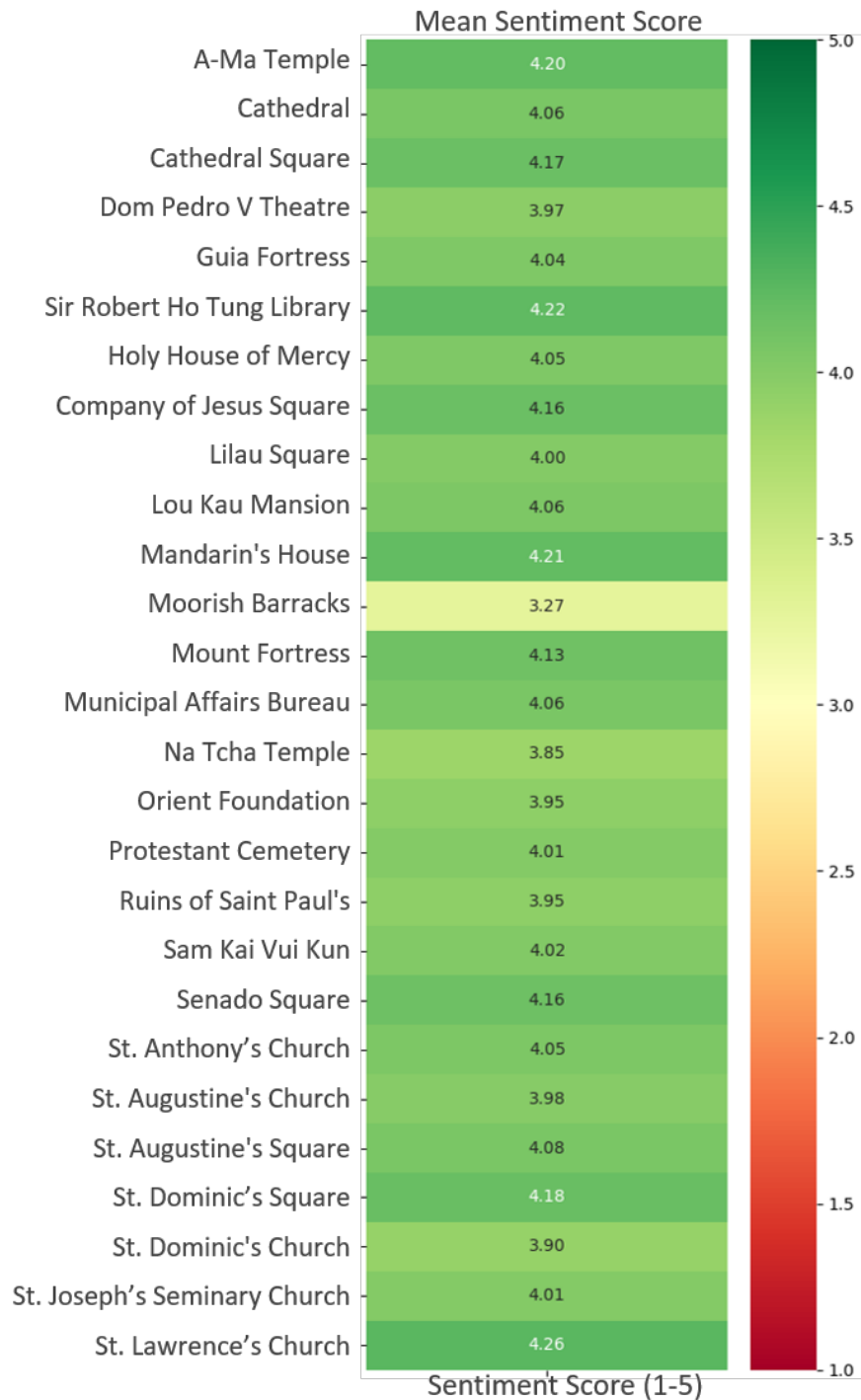


Figure 5.13.: Mean sentiment score (1–5) by heritage component.

5.3.2. Keyterm Profiles and Comparisons

System-wide keyterm extraction provides a complementary view of what visitors highlight in their written descriptions across the heritage system. Figure 5.14 offers a qualitative overview of the vocabulary space. The word cloud is used here as an

walk through), alongside recurring component labels.

Across platforms (Figure 5.16), Google Maps exhibits the most consistently high and diverse keyword importance values across the list, suggesting a comparatively heterogeneous vocabulary that spans component names, architectural descriptors, and visit-practice terms. But perhaps it’s because it contains the largest number of reviews in three languages. By contrast, Ctrip displays a more selective profile with pronounced spikes in a smaller number of terms (e.g., *entrance* and *mount song*), while TripAdvisor concentrates more strongly on a subset of architectural and building-related terms (e.g., *architectural*, *mansion*, and several church- or order-linked names). Overall, the heatmaps indicate that language and platform shape which components and descriptive lenses are highlighted, even though a shared heritage lexicon is present across the system.

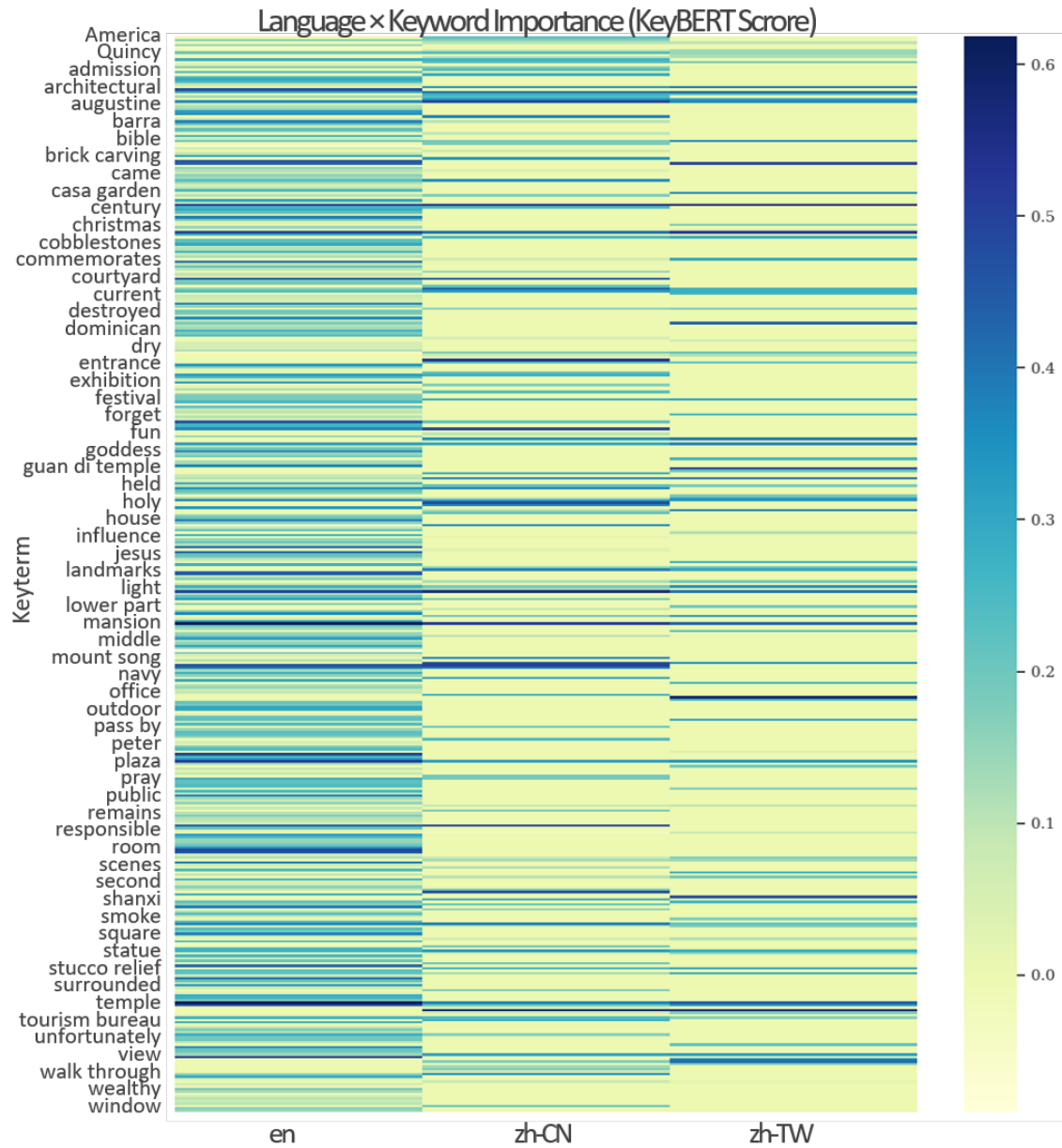


Figure 5.15.: System-wide keyterm importance by language (KeyBERT scores). Darker cells indicate higher importance for a given keyterm within each language subset. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.

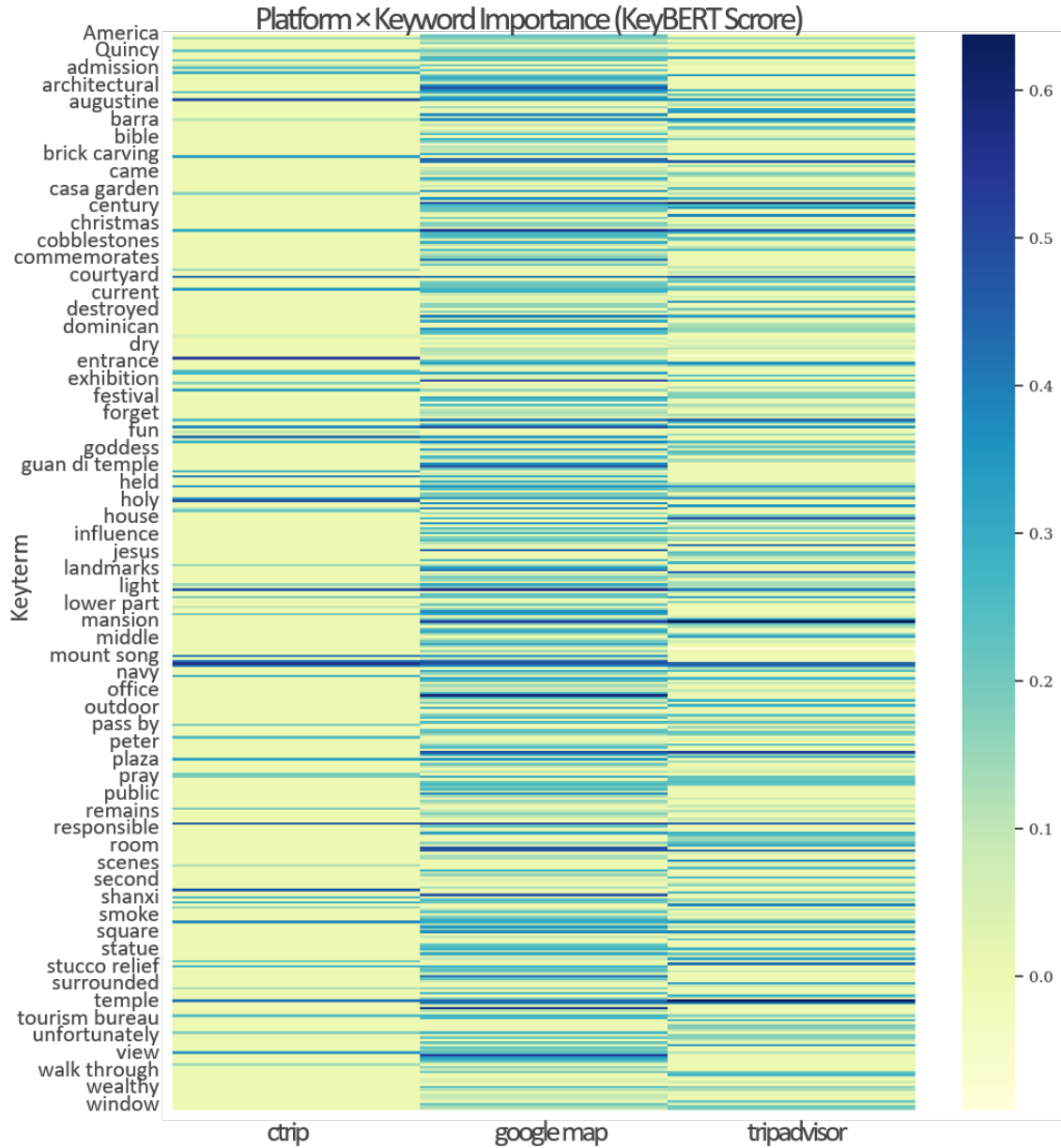


Figure 5.16.: System-wide keyterm importance by platform (KeyBERT scores). Darker cells indicate higher importance for a given keyterm within each platform subset.

To move beyond visual inspection of the system-wide heatmaps, keyword rankings with pairwise Mann–Whitney U tests are used again. Because the system-wide vocabulary is large, the analysis was restricted to the top 20 most divergent keywords (based on the KeyBERT-derived divergence score), which makes the comparison interpretable while still targeting the terms where subsets differ most.

Figure 5.17 shows that rank divergence is concentrated between the two Chinese varieties: the zh-CN vs zh-TW comparison is significant (***, $p < 0.001$), whereas

English vs Chinese (en vs zh-CN; en vs zh-TW) is not significant in this top-20 set. The strongest shifts occur in a mix of built-form descriptors and component-linked terms (e.g., *building*, *architectural*, *ruins*, and terms such as *Shakespeare*, *old house*), indicating that Simplified and Traditional Chinese reviews do not prioritise the same place labels and descriptive lenses to the same degree, even within the same heritage system.

Across platforms (Figure 5.18), only the Ctrip vs TripAdvisor comparison is significant (***, $p < 0.001$). The heatmap shows a systematic separation in rank positions across many of the selected keywords. In particular, several component- or institution-linked terms (e.g., *old house*, *tombstone*, *villa*, *guan di temple*, *rosary*, and *shakespeare*) exhibit large rank shifts between the two platforms, and multiple built-form descriptors (e.g., *architectural*, *building*, *house*, and *mansion*) also contribute to the divergence. By contrast, the other platform pairs (Ctrip vs Google Maps; Google Maps vs TripAdvisor) are not significant in this top-20 set, suggesting that Google Maps and TripAdvisor are closer in their rank distributions than either is to Ctrip in this filtered comparison. Overall, the significant Ctrip–TripAdvisor contrast indicates platform-specific prioritisation of which named components and descriptive lenses are mentioned.

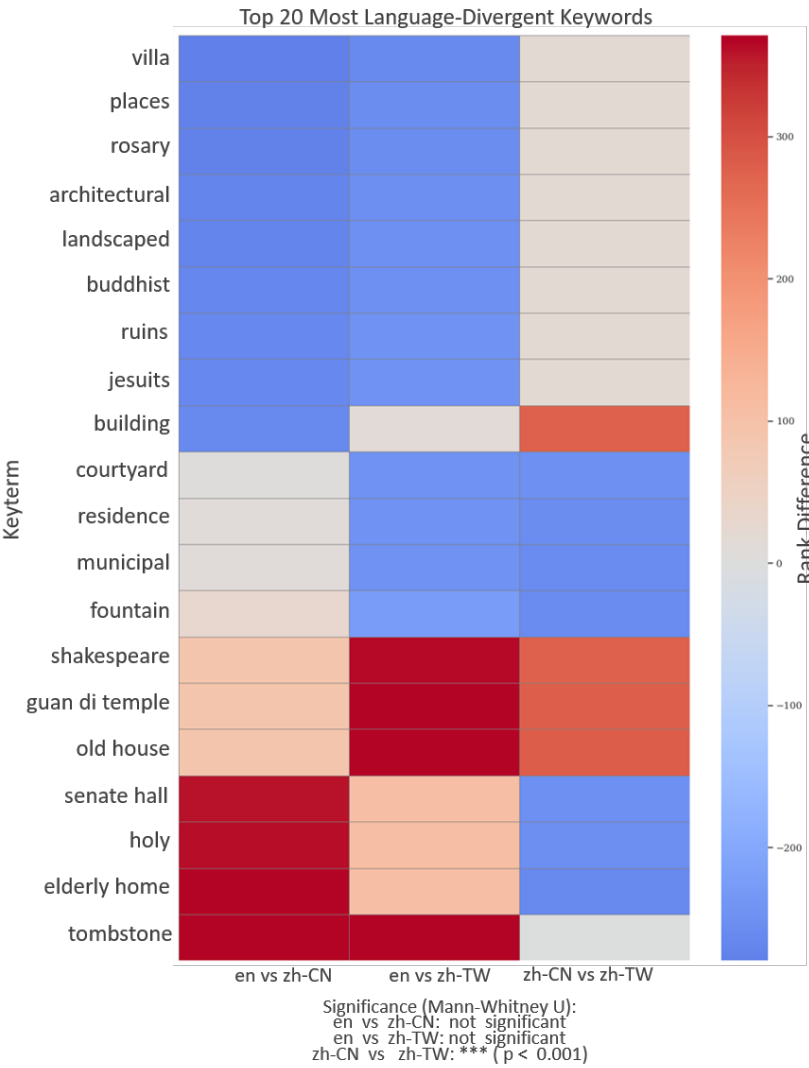


Figure 5.17.: Top 20 most language-divergent keywords. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.

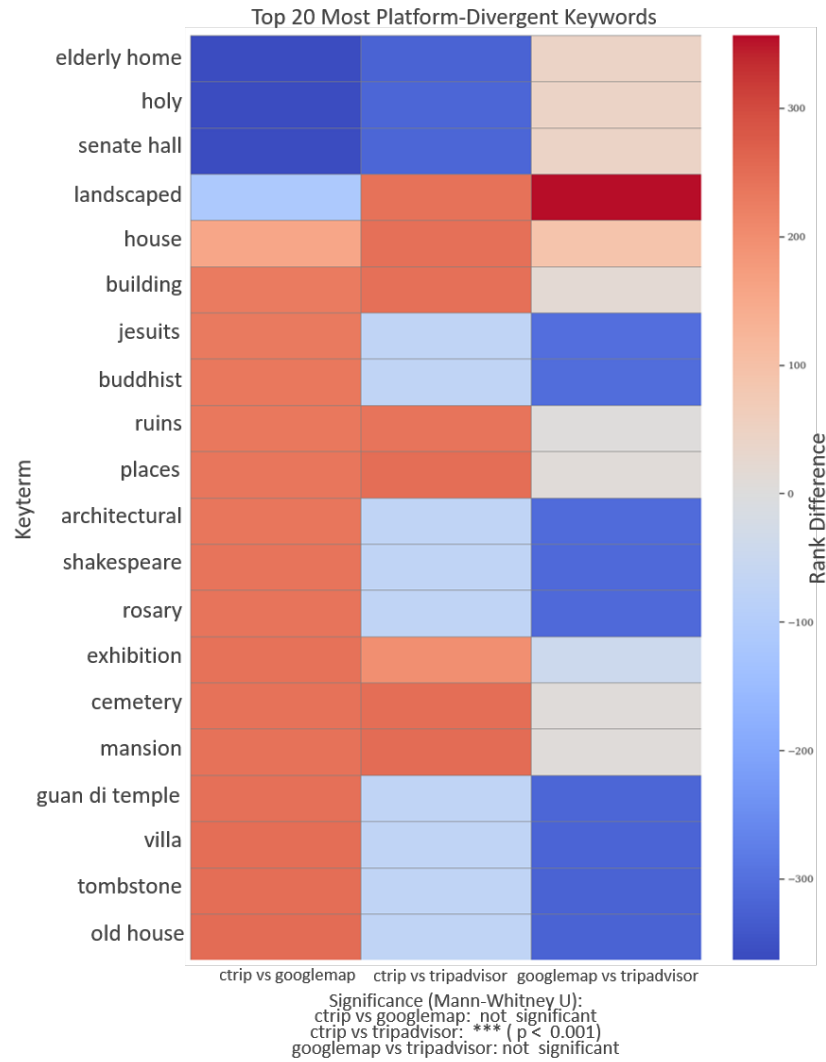


Figure 5.18.: Top 20 most platform-divergent keywords.

5.3.3. OUV Alignment: Global Patterns

This section summarises system-wide patterns from the generic OUV classification and Macao-Specific OUV Alignment outputs. The subset of sentences retained as OUV-relevant and assigned to one of the four UNESCO criteria used for the Historic Centre of Macao. For interpretability, the criteria are discussed using dominant analytical lenses: Criterion (ii) (*cultural–intellectual exchange mechanisms*), Criterion (iii) (*contact zone & cultural fusion*), Criterion (iv) (*material urban fabric & architectural ensembles*), and Criterion (vi) (*circulation of ideas & socio-political change*).

At the aggregate level, the same as anchor case, the criterion distribution in Figure 5.19 is strongly biased to Criterion (iv) (*material urban fabric & architectural*

ensembles). In other words, when reviewers express OUV-aligned content, they most often talk about the tangible built environment: streetscapes, church ensembles, and architectural features, rather than abstract historical interpretation alone. Criterion (iii) (*contact zone & cultural fusion*) forms a consistent secondary share, reflecting recurrent framing of Macao as a long-term contact zone, while Criterion (ii) (*cultural-intellectual exchange mechanisms*) and Criterion (vi) (*circulation of ideas & socio-political change*) still account for comparatively small proportions of aligned sentences in the system-wide total.

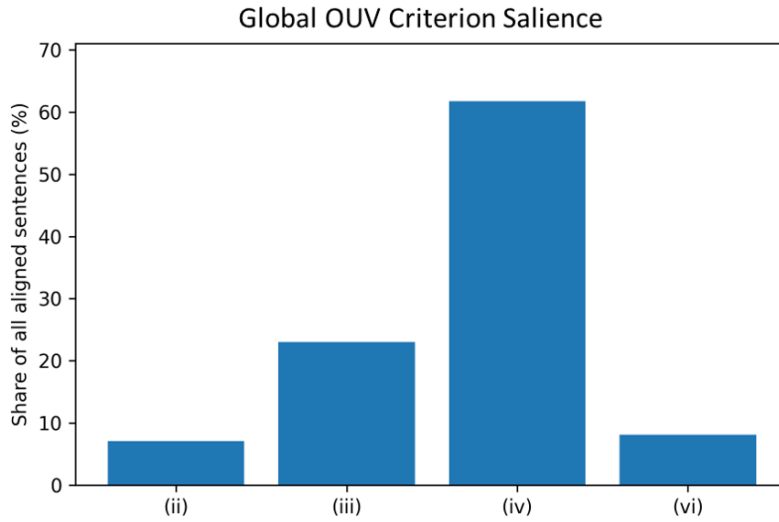


Figure 5.19.: System-wide criterion saliency from Macao-Specific OUV Alignment across all components in the Historic Centre of Macao.

Component-level profiles show that this Criterion (iv) (*material urban fabric & architectural ensembles*) dominance is broadly stable across the heritage system (Figure 5.20). Most components exhibit their highest shares under Criterion (iv) (*material urban fabric & architectural ensembles*), indicating that the alignment signal is driven by material descriptions across diverse site types (churches, squares, temples, and civic buildings). The heatmap shows meaningful variation in the balance between Criterion (iii) (*contact zone & cultural fusion*) and the other criteria for particular components. For instance, components with clearer institutional, cultural, or programme-related associations (e.g., theatres and libraries) display a comparatively stronger presence of Criterion (iii) (*contact zone & cultural fusion*) and Criterion (vi) (*circulation of ideas & socio-political change*), while several residential or civic components show a more noticeable Criterion (iii) (*contact zone & cultural fusion*) share alongside Criterion (iv) (*material urban fabric & architectural ensembles*).

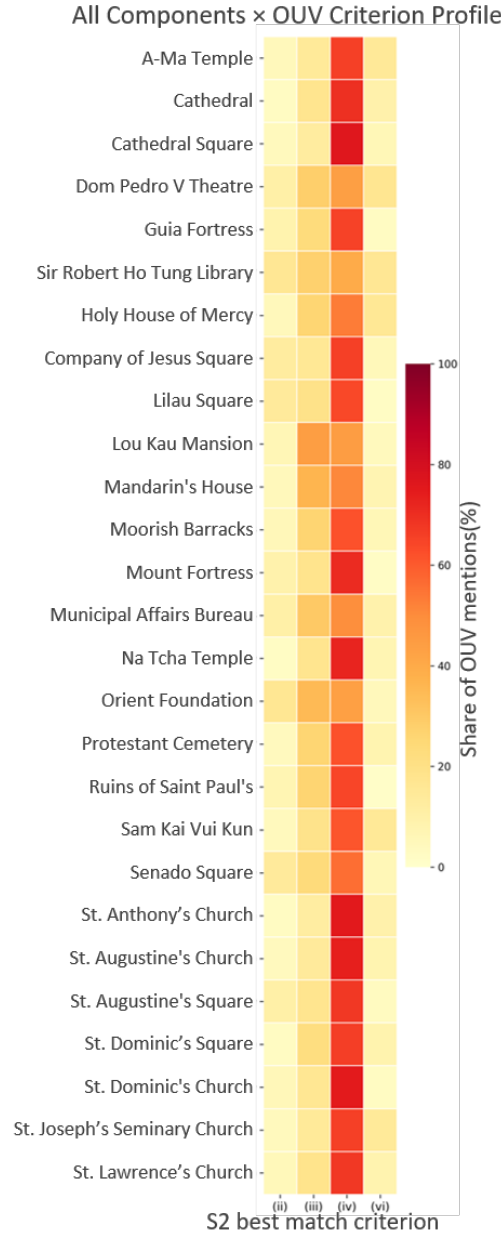


Figure 5.20.: Component-level criterion from Macao-Specific OUV Alignment. Each cell highlights how criterion emphasis varies across components.

When disaggregated by language and platform (Figure 5.21), the criterion profile is structurally stable across subsets. In all language groups (English, Simplified Chinese, Traditional Chinese) and on all platforms (Ctrip, Google Maps, TripAdvisor), Criterion (iv) (*material urban fabric & architectural ensembles*) remains the dominant category, with Criterion (iii) (*contact zone & cultural fusion*) consistently second. The main differences across subsets are confined to relative magnitudes: small shifts appear in the minor categories Criterion (ii) (*cultural-intellectual exchange mechanisms*) and Criterion (vi) (*circulation of ideas & socio-political change*), and in how strongly Criterion (iv) (*material urban fabric & architectural ensem-*

bles) outweighs the remainder. Overall, the results suggest that language and platform modulate emphasis at the margins, but do not alter the dominant OUV lens expressed in the aligned sentences.

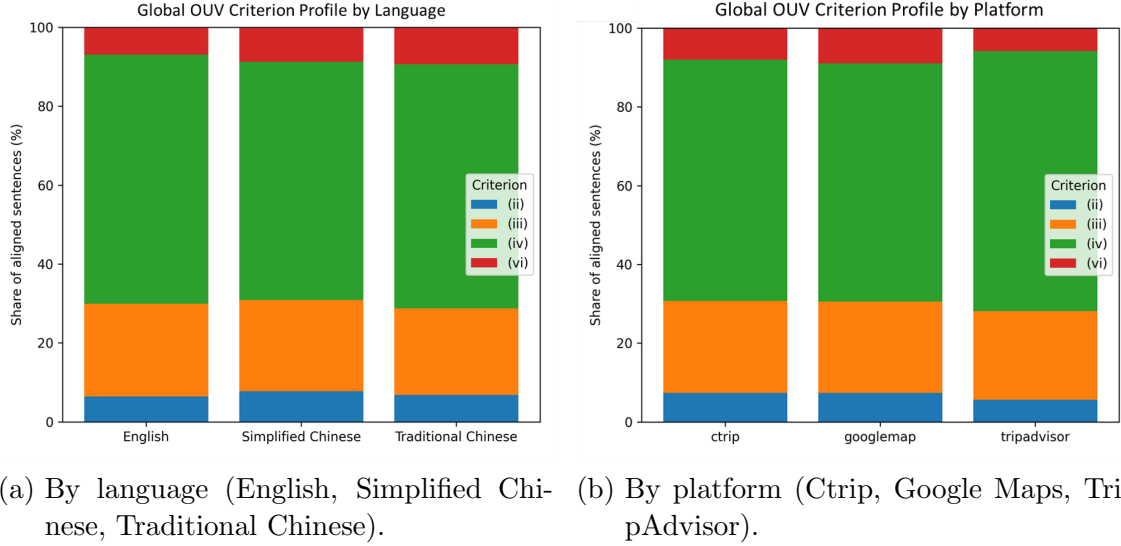


Figure 5.21.: System-wide criterion profiles from Macao-Specific OUV Alignment, disaggregated by language and by platform. Stacked bars show the share of aligned sentences assigned to each criterion within each subset.

To illustrate how the two-stage pipeline reshapes criterion assignment, Figure 5.22(a) contrasts the criterion profiles produced by the Generic OUV Classification and the Macao-Specific OUV Alignment. The generic model yields a highly concentrated distribution dominated by Criterion (ii) (*cultural–intellectual exchange mechanisms*), with a much smaller share attributed to Criterion (iv) (*material urban fabric & architectural ensembles*) and only marginal shares for Criterion (iii) (*contact zone & cultural fusion*) and Criterion (vi) (*circulation of ideas & socio-political change*). After applying the Macao-Specific OUV Alignment, the profile shifts decisively: Criterion (iv) (*material urban fabric & architectural ensembles*) becomes the dominant category, Criterion (iii) (*contact zone & cultural fusion*) rises to a substantial secondary share, and Criterion (vi) (*circulation of ideas & socio-political change*) also increases, while Criterion (ii) (*cultural–intellectual exchange mechanisms*) is reduced to a minor proportion. In practical terms, the Macao-specific interpretation re-centres the aligned discourse on the built urban fabric and architectural ensembles, with contact-zone framing and broader idea-circulation appearing as secondary lenses.

Figure 5.22(b) shows the corresponding transition matrix (row-normalised by Generic OUV Classification labels). Across all four generic labels, most sentences are reassigned to Criterion (iv) (*material urban fabric & architectural ensembles*) in the Macao-Specific OUV Alignment: for example, generic Criterion (ii) (*cultural–intellectual*

exchange mechanisms) sentences map predominantly to Criterion (iv) (*material urban fabric & architectural ensembles*)(61.9%) and Criterion (iii) (*contact zone & cultural fusion*)(23.5%); generic Criterion (iii) (*contact zone & cultural fusion*) sentences map mainly to Criterion (iv) (*material urban fabric & architectural ensembles*)(71.2%) and Criterion (vi) (*circulation of ideas & socio-political change*)(22.0%); generic Criterion (iv) (*material urban fabric & architectural ensembles*) sentences still map mostly to Criterion (iv) (*material urban fabric & architectural ensembles*)(60.1%) but also shift toward Criterion (iii) (*contact zone & cultural fusion*)(22.2%); and generic Criterion (vi) (*circulation of ideas & socio-political change*) sentences overwhelmingly map to Criterion (iv) (*material urban fabric & architectural ensembles*)(89.0%). This pattern indicates that the Macao-specific alignment reinterprets generic OUV-related statements through the Macao criteria, concentrating the final aligned signal in Criterion (iv) (*material urban fabric & architectural ensembles*) and, secondarily, Criterion (iii) (*contact zone & cultural fusion*).

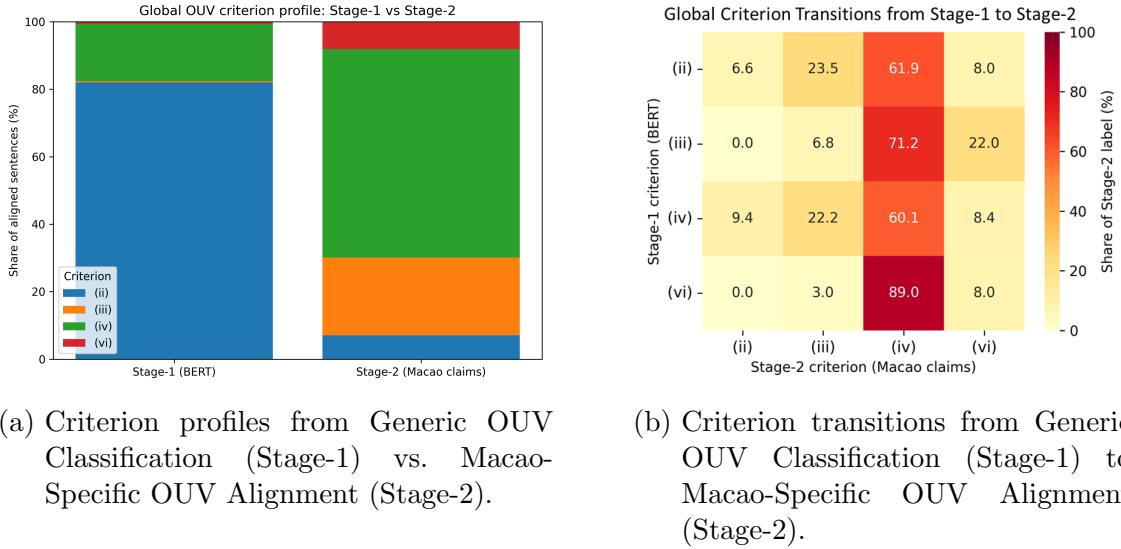


Figure 5.22.: System-wide comparison of Generic OUV Classification and Macao-Specific OUV Alignment.

5.4. System-wide Visual Patterns

5.4.1. System-wide Visual Theme Model

As described in Section 4.5, visual themes were learned on the full photo collection pooled across all heritage components to capture system-wide regularities in visitor imagery. Each image was embedded into a fixed-length visual feature vector using a pretrained vision encoder, and K-means clustering with $K = 8$ was applied in

this embedding space to assign every image to a theme. Cluster labels were then interpreted and named through manual inspection of representative samples. For interpretability, a t-SNE projection of a 600-image subset is shown as a qualitative diagnostic. It provides an intuitive view where visually similar images tend to lie closer, but it is not used as evidence that clusters are strictly separable.

Figure 5.23 summarises the system-wide distribution of visual themes across all images. The distribution is uneven, with *Interiors & Decorative Details* and *Green Spaces & Parks* emerging as the most frequent themes, followed by *Facades* and *Buildings & Streetscapes*. Themes related to religious interiors, folk shrines, and the Guia Fortress appear less frequently, indicating that a small number of visually dominant motifs account for a substantial share of visitor imagery at the system level.

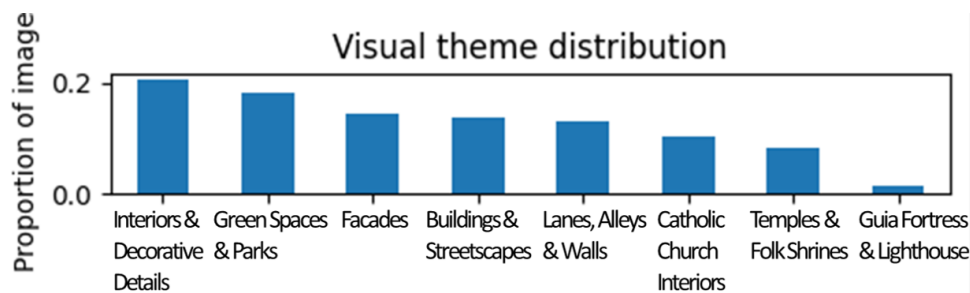


Figure 5.23.: System-wide distribution of visual themes across all images, ordered by decreasing proportion.

5.4.2. Visual Theme Definitions and Exemplars

To make the eight visual themes interpretable, each cluster was defined using a short operational description and illustrated with a small set of representative images (Figure 5.24). Exemplars are shown as qualitative anchors: they clarify what the model groups together, but they are not intended to summarise within-theme variability or to serve as a quantitative evaluation.

Temples & Folk Shrines Images centred on Chinese religious architecture and ritual spaces (temple gates, incense burners, altars, deity statues), often with courtyard context.

Buildings & Streetscapes Wider views of the built environment that situate heritage structures in their urban setting (street frontage, plazas, approach routes, surrounding blocks).

Lanes, Alleys & Walls Narrow passages, stone walls, steps, side streets, and tran-

sitional spaces that emphasise texture and enclosure rather than landmark façades.

Interiors & Decorative Details Close-range interior scenes and material details (ceilings, tiles, reliefs, inscriptions, ornamentation), typically framed at object level.

Guia Fortress & Lighthouse Images dominated by the Guia lighthouse and fortress complex and hilltop viewpoints, including exterior shots and panoramic over-looks.

Green Spaces & Parks Vegetation-forward scenes (gardens, trees, lawns, landscaped paths) where greenery is the primary visual subject.

Catholic Church Interiors Nave interiors and liturgical elements inside churches (pews, altars, stained glass), distinct from exterior streetscape views.

Facades Tight exterior shots focusing on a single building face (doorways, windows, elevations), typically excluding the broader plaza and streetscape context.



(a) C0: Temples & Folk Shrines (b) C1: Buildings & Streetscapes (c) C2: Lanes, Alleys & Walls (d) C3: Interiors & Decorative Details



(e) C4: Guia Fortress & Lighthouse (f) C5: Green Spaces & Parks (g) C6: Catholic Church Interiors (h) C7: Facades

Figure 5.24.: Visual theme exemplars for the clustering solution ($K = 8$).

5.4.3. Embedding-space Structure of Themes

Figure 5.25 shows a t-SNE projection of 600 sampled images in the learned embedding space, coloured by the $K = 8$ K-means theme assignment. The themes are only partly separable. A few categories form relatively compact regions, most notably *Catholic Church Interiors* and *Guia Fortress & Lighthouse*, suggesting that these motifs have distinctive and consistently captured visual cues. In contrast, outdoor built-environment themes (especially *Buildings & Streetscapes*, *Facades*, and *Lanes, Alleys & Walls*) appear close to each other and overlap, reflecting shared architectural textures and a viewpoint continuum from wide contextual shots to facade close-ups and narrow passages.

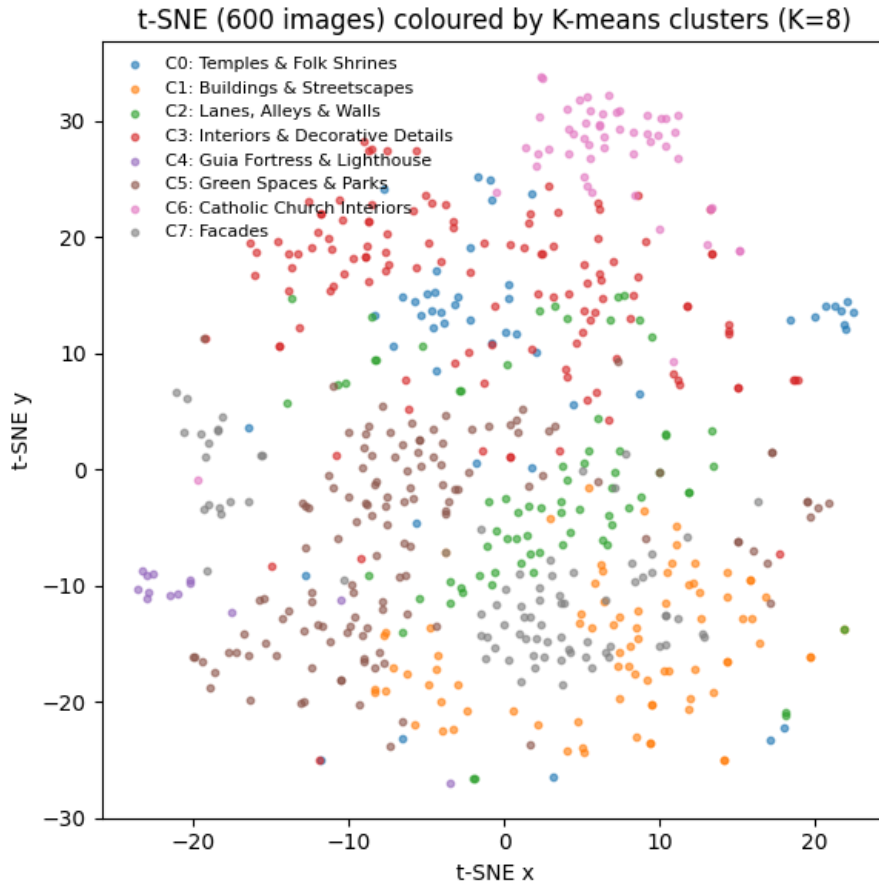


Figure 5.25.: **Embedding-space structure of system-wide visual themes.** t-SNE projection of 600 randomly sampled images in the learned embedding space, coloured by K-means theme assignment ($K = 8$).

Figure 5.26 provides a qualitative check using the same sample images. The presence of visually similar architectural photos across neighbouring theme colours is consistent with the overlap in Figure 5.25: many visitor images contain multiple cues (e.g., a facade together with its plaza), while K-means assigns a single dominant label. Accordingly, the eight themes are treated as dominant visual emphases,

and overlap is interpreted as gradual transitions in photographic composition.

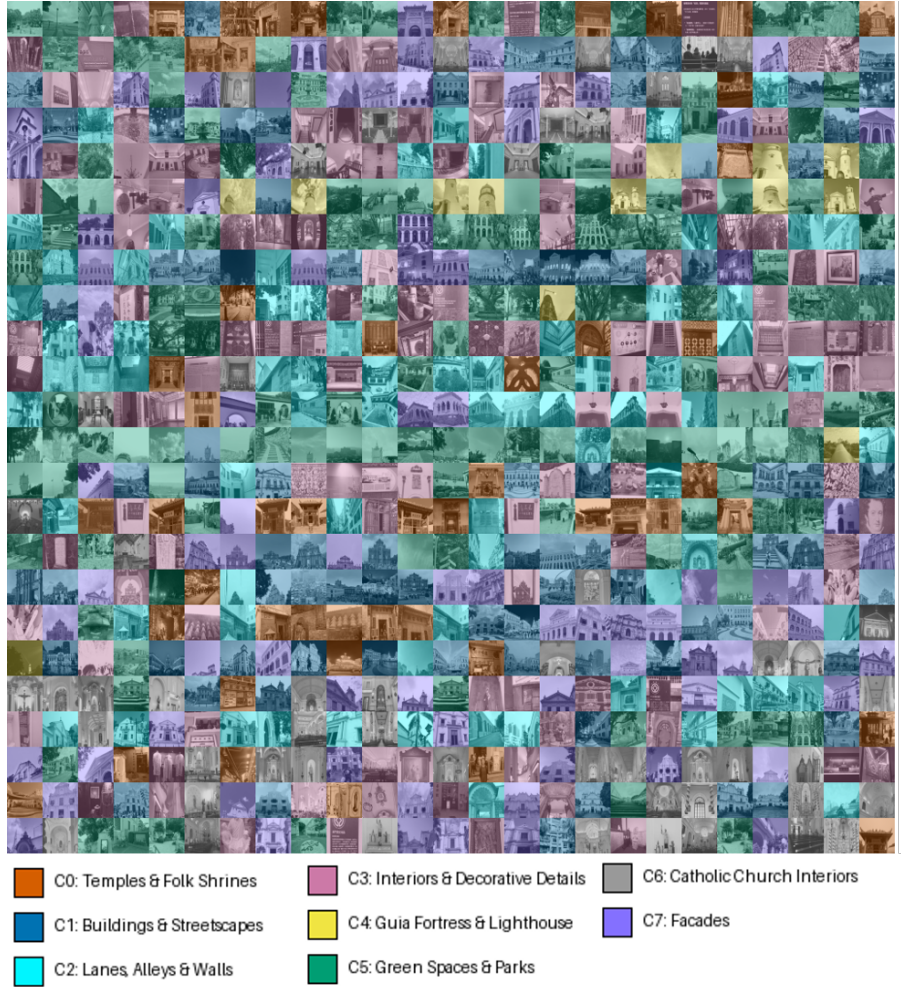


Figure 5.26.: **Qualitative exemplars across the embedding space.** Thumbnail wall of the same 600 sampled images, tinted by their assigned visual theme.

5.4.4. Language-sensitive Preferences

Figure 5.27 presents POI-normalised visual theme preferences across all heritage components. At this aggregate scale, visual theme distributions are highly similar across English, Simplified Chinese, and Traditional Chinese images, indicating strong convergence in visual representation patterns irrespective of language. Language-related differences are limited to minor shifts in relative emphasis, such as a slightly reduced presence of *Guia Fortress & Lighthouse* imagery in Simplified Chinese images and a marginally higher emphasis on interior details in Traditional Chinese images. Taken together, these results suggest that language plays a secondary role in shaping system-wide visual patterns, with dominant visual representations primarily driven by the characteristics of the heritage components themselves.

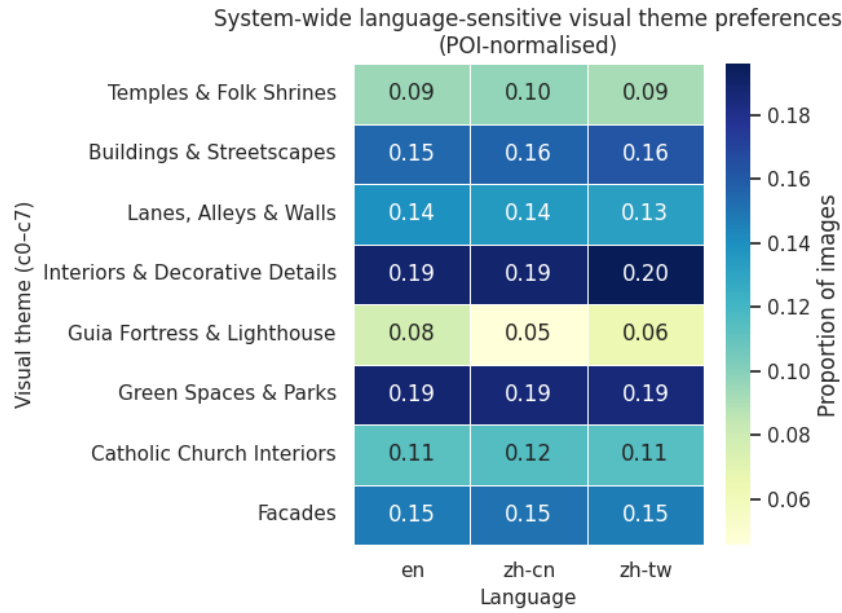


Figure 5.27.: System-wide language-sensitive visual theme preferences across all heritage components. en = English, zh-CN = Simplified Chinese, zh-TW = Traditional Chinese.

5.4.5. Component Visual Theme Profiles

Figure 5.28 shows, for each component, the percentage of images assigned to each of the eight visual themes (rows sum to 100%). This makes it easy to see which themes dominate a component’s photo set and which components share similar visual signatures.

As shown in subsection 5.2.4, Ruins of St. Paul’s is primarily characterised by *Buildings & Streetscapes* and *Facades*. Several other components display similarly distinctive peaks that align with their most recognisable motifs, for example Guia Fortress with *Guia Fortress & Lighthouse*, and selected temple sites (e.g., Na Tcha Temple and Sam Kai Vui Kun) with *Temples & Folk Shrines*.

At the same time, the heatmap highlights cross-component regularities. Many Catholic sites show comparatively higher shares of *Catholic Church Interiors*, suggesting that interior photography is a common practice across multiple churches rather than a single-site exception. Civic squares and open nodes tend to combine *Buildings & Streetscapes* with *Facades*, while mansion-type heritage locations show relatively stronger *Interiors & Decorative Details*. Overall, the component profiles indicate that the visual themes are interpretable at site level and reveal both distinctive landmarks and shared photographic routines across the heritage system.

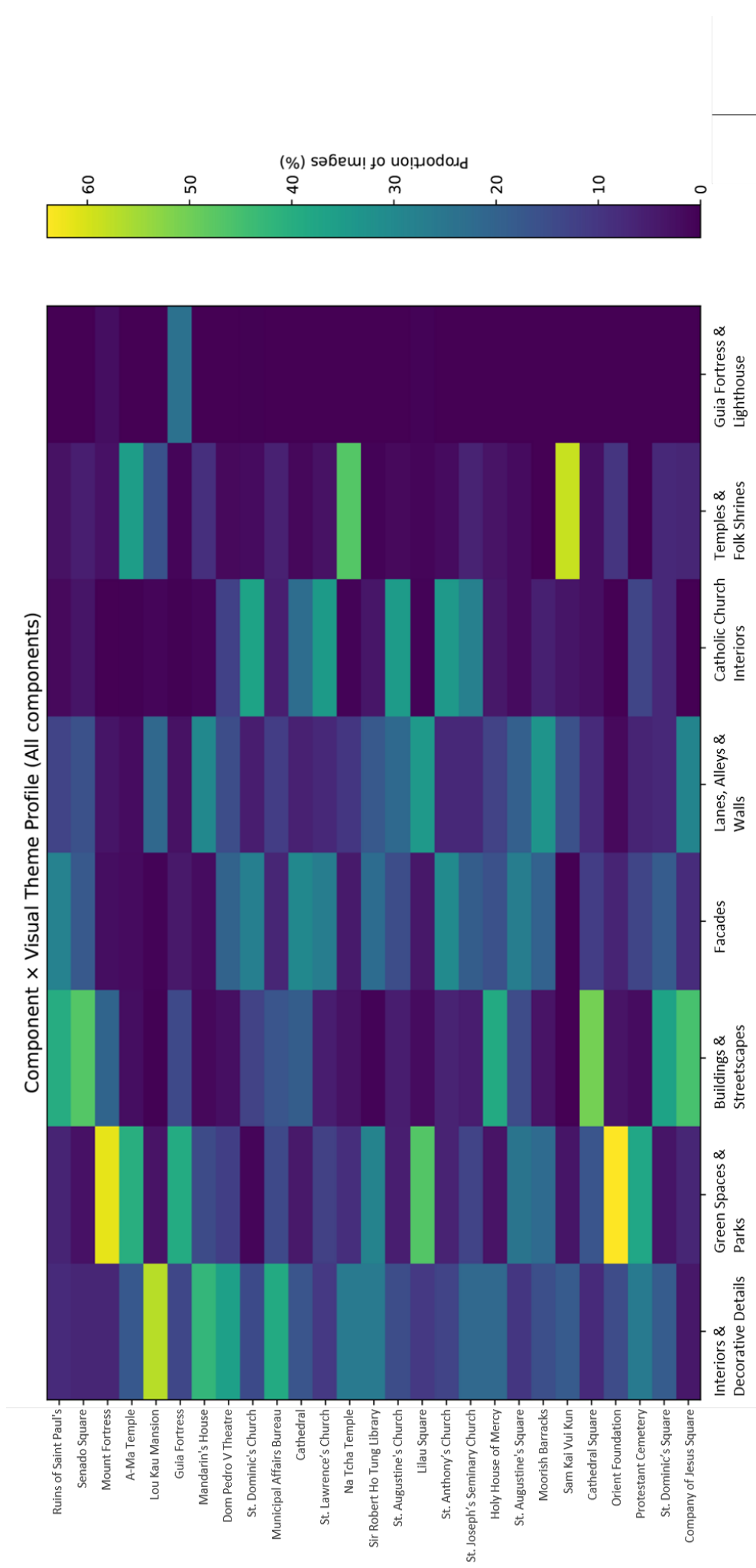


Figure 5.28.: Component $\times$ visual theme profile across all heritage components. Cells report the proportion of images assigned to each visual theme within a component (row-normalised). Higher values indicate that a theme is more prevalent in that component's photo set.

5.5. Cross-modal Synthesis

5.5.1. Component-level alignment across text and images

This section links textual and visual signals at the component level by comparing aggregated visual theme profiles with component-level textual profiles derived from OUV alignment, keyterm signals, and sentiment. The focus is not on individual sentences or images, but on whether components dominated by particular visual themes also tend to foreground specific value framings and lexical cues in the associated review discourse. For transparency, a component-level synthesis dashboard reporting sentence volume, mean sentiment, top keyterms, Stage-2 OUV criterion composition (II/III/IV/VI), and dominant visual theme for all 27 components is provided in Appendix A.4.

Figure 5.29 presents Spearman correlations between visual theme proportions and OUV claim proportions across components. Several moderate associations suggest non-random alignment between what is photographed and which dimensions of OUV are emphasised in the text. For example, *Buildings & Streetscapes* show a positive association with Criterion (iv) (*material urban fabric & architectural ensembles*), while *Catholic Church Interiors* and *Interiors & Decorative Details* are positively associated with Criterion (vi) (*circulation of ideas & socio-political change*). In contrast, *Green Spaces & Parks* and *Interiors & Decorative Details* exhibit negative associations with Criterion (iv) (*material urban fabric & architectural ensembles*), indicating that components dominated by interior or landscape imagery tend to foreground different aspects of OUV in textual narratives.

While the OUV correlations capture value-oriented framing, the review discourse also contains recurrent lexical cues that help contextualise these alignments. Figure 5.30 summarises the strongest visual-lexical associations by reporting the top-10 keyterms selected by maximum absolute Spearman correlation across themes. The resulting pattern is not diffuse: several visual themes show coherent lexical companions, indicating structured alignment between dominant image motifs and the vocabulary used to describe components. In particular, religious framings separate clearly. "Church" aligns positively with *Catholic Church Interiors* and *Facades*, whereas *Temples & Folk Shrines* aligns strongly with "Temple", "China", "Incense", and exhibits the opposite pattern for built-form and colonial vocabulary (e.g., "Archway", "Colonial"). Landmark-specific pairing is also evident: "Fortress" peaks for *Guia Fortress & Lighthouse*. Finally, spatial terms track scene orientation: "courtyard" is most strongly associated with *Interiors & Decorative Details*

and shows weaker or negative associations with exterior-dominant themes. Overall, the heatmap supports the view that components characterised by particular visual emphases are accompanied by systematically different lexical signals, consistent with cross-modal structural alignment.

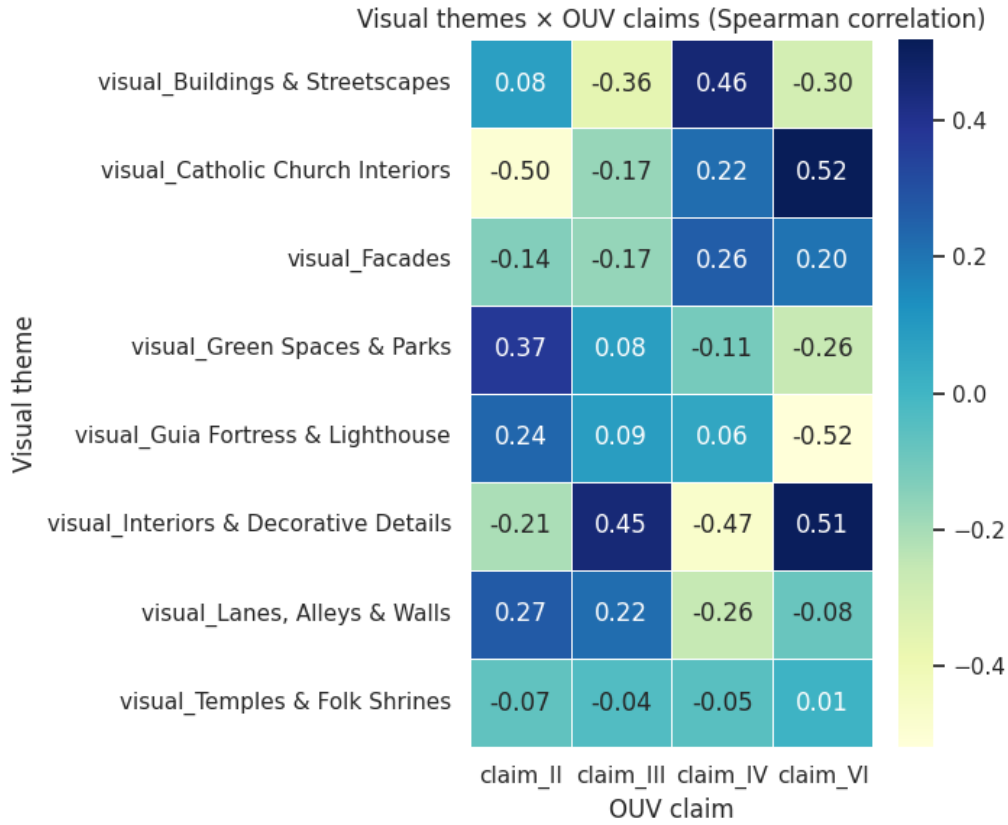


Figure 5.29.: Spearman correlations between visual theme proportions and OUV claim proportions across heritage components, indicating rank-based structural alignment between visual and textual representations.

To separate thematic alignment from evaluative tone, Figure 5.31 shows average sentiment scores by OUV claim. Sentiment is consistently positive across criteria, with only modest variation, suggesting that differences observed in the OUV and keyterm correlation patterns are primarily driven by differences in narrative focus rather than by strong differences in affective polarity. Criterion (iii) (*contact zone & cultural fusion*) and Criterion (iv) (*material urban fabric & architectural ensembles*) exhibit slightly higher average sentiment.

Figure 5.32 further relates visual themes to sentiment at the component level. *Buildings & Streetscapes* display the strongest positive association with sentiment, followed by *Facades* and *Guia Fortress & Lighthouse*. In contrast, *Lanes, Alleys & Walls* and *Green Spaces & Parks* are weakly negatively associated with sentiment, suggesting that components dominated by these visual themes are described in a

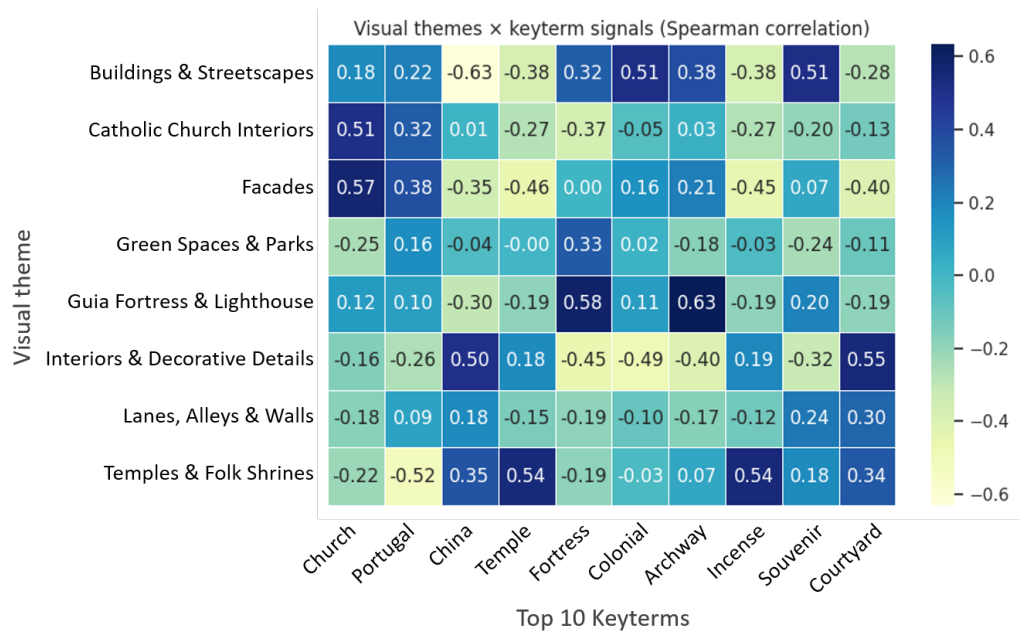


Figure 5.30.: Spearman correlations between visual theme proportions and top-10 keyterms (selected by maximum absolute association across themes), indicating how lexical framing co-varies with dominant visual emphases across components.

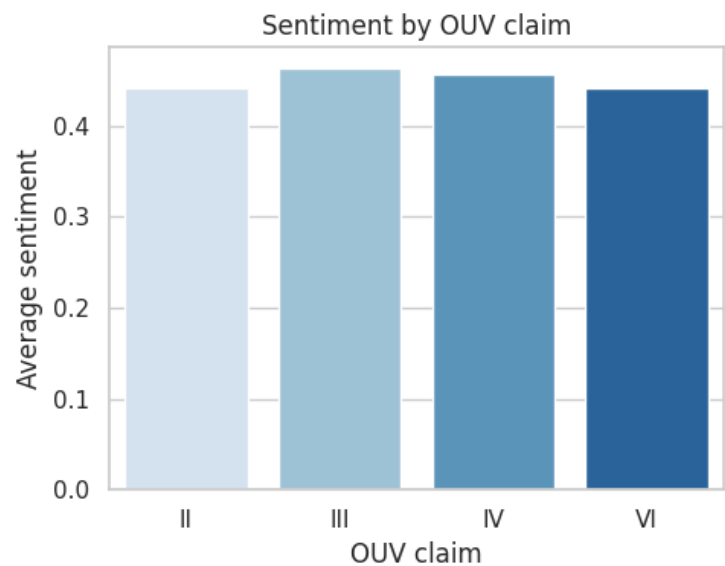


Figure 5.31.: Average sentiment scores of OUV-aligned sentences by claim, showing consistently positive sentiment across claims with limited variation.

more neutral or mixed evaluative tone. Together, these results indicate that visual dominance is linked not only to which OUV narratives and lexical frames are expressed, but also to how positively those narratives are framed.

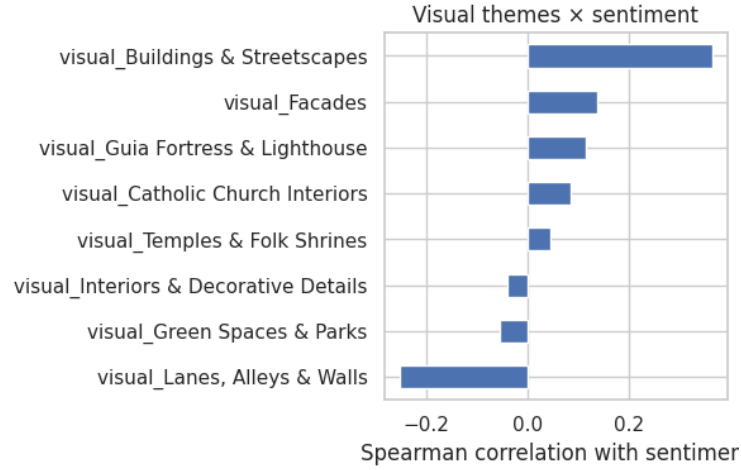


Figure 5.32.: Spearman correlations between visual theme dominance and average sentiment at the component level, indicating how visual emphasis relates to evaluative framing.

Overall, the component-level comparison indicates that textual and visual representations are not independent. Instead, certain visual themes consistently co-occur with specific OUV claims, keyterm signals, and sentiment profiles, supporting a structured alignment between what visitors photograph and how they articulate heritage value.

5.5.2. Summary of Cross-Modal Findings

This chapter has presented results from a unified analysis pipeline that links textual and visual representations at the level of heritage components. On the textual side, visitor sentences were aligned to a small set of Macao-relevant OUV criteria and summarised using claim shares, keyterm signals, and sentiment scores. On the visual side, images were grouped into a limited number of visual themes, producing component-level visual profiles.

When both modalities are considered together, the results show that textual and visual representations are related in systematic ways. Components with a strong emphasis on exterior architectural imagery are more strongly associated with Criterion (iv) (*material urban fabric & architectural ensembles*)-related narratives, while components dominated by interior imagery are more strongly associated with Criterion (vi) (*circulation of ideas & socio-political change*). *Interiors & Decorative*

Details also shows a positive association with Criterion (iii) (*contact zone & cultural fusion*) and a negative association with Criterion (iv) (*material urban fabric & architectural ensembles*). Beyond OUV alignment, the keyterm correlations indicate coherent visual–lexical pairings rather than diffuse associations: Catholic interior and facade-dominant components align with church-related vocabulary, whereas temple-dominant components align with terms such as temple/China/incense, and landmark-specific imagery (e.g., *Guia Fortress & Lighthouse*) aligns with fortress-related terms. Sentiment scores are positive across all OUV criteria and show only small differences, while correlations between visual themes and sentiment indicate that components dominated by landmark and streetscape imagery tend to be framed more positively than those dominated by transitional urban elements.

These component-level alignments remain broadly consistent when the textual data are examined by platform and language, with differences appearing mainly as modest shifts in emphasis rather than changes in the overall pattern. Finally, the anchor case of the Ruins of St. Paul’s illustrates how these system-wide patterns appear at a single heritage component through its dominant visual themes, associated OUV narratives, key lexical cues, and sentiment framing.

Overall, the results offer an integrated view of how the Historic Centre of Macao is represented in public user-generated content. Across both textual and visual modalities, a small set of recurring themes and values consistently emerges. Reviews foreground architectural features, interior spaces, and cultural or symbolic meanings, and these framings are mirrored in what visitors choose to photograph and in the vocabulary they use to describe components. In this sense, the findings demonstrate a correspondence between user-generated representations and key dimensions of Outstanding Universal Value defined by UNESCO. The following chapter builds on these empirical results to discuss how and why these patterns emerge, and what they imply for understanding heritage representation in digital discourse.

6. Discussion

Digital traces such as reviews and photographs are increasingly used to study how places are perceived and represented, complementing more traditional elicitation approaches such as surveys, interviews, and participatory methods. In the place and landscape literature, computational work has shown that first-person texts can be analysed to recover recurring perceptual descriptors and to produce interpretable “character assessments” that link language use back to physical settings. This line of research treats public text not merely as commentary, but as evidence of how people categorise and make salient features of environments, and it motivates the use of large-scale, naturally occurring descriptions as a window into place meanings (Koblet and Purves, 2020; Wartmann et al., 2018; Purves et al., 2023b).

Within tourism and heritage contexts, these approaches become particularly relevant because the “image” of a destination is continuously reproduced through platform-specific narratives and visuals, rather than only through official interpretation (Gartner, 1994; Van Dijck et al., 2018). The Historic Centre of Macao provides a clear case where formal OUV statements coexist with dense everyday representations circulating across languages and platforms (Jokilehto, 2008). The results in this thesis therefore contribute not only to UGC-based heritage analytics, but more broadly to research on how place meanings are expressed through language and imagery under contemporary platform conditions, and how these expressions align with formal value frameworks (Zeng and Gerritsen, 2014; Bai et al., 2021).

In the digital age, these expressions increasingly leave traces in online environments, where people describe, rate, and visually curate places through platform-specific conventions (Litvin et al., 2008; Xiang and Gretzel, 2010). Against this broad backdrop, this chapter discusses the main findings of the thesis in relation to the research questions and the wider literature on public perception of urban heritage, tourism destination image, and user-generated content (UGC).

The preceding chapters have shown how multilingual reviews and user-uploaded images from Google Maps, TripAdvisor, and Ctrip can be organised in a place-based framework for the Historic Centre of Macao, and how natural language processing

and computer vision methods can be used to derive indicators of sentiment, themes, and value alignment. Here, the focus shifts from technical details to interpretation: what the results suggest about how visitors perceive the Historic Centre, how these perceptions relate to the Outstanding Universal Value (OUV) of the property, and what this might mean for World Heritage management.

The chapter is structured around the four research questions introduced in Chapter 1. The first part revisits how visitors describe and emotionally evaluate the Historic Centre of Macao in their multilingual online reviews, and how these narratives relate to everyday experiences of the site. The second part discusses the relationship between these public narratives and the OUV criteria adopted for the property, drawing on debates about the gap between expert and public understandings of heritage (Smith, 2006; Labadi, 2013). The third part turns to user-generated images and examines which system-wide visual themes and language-sensitive preferences can be observed in the photographic representations of the heritage components. The fourth part then considers how textual and visual UGC complement or diverge from each other, connecting the empirical findings to concepts. This synthesis follows the cross-modal integration strategy summarised in Figure 4.6, which aligns textual, OUV, and visual outputs at the level of heritage components. The chapter concludes with methodological reflections and outlines implications for heritage management and future research.

6.1. Revisiting the Research Questions

The first research question asked how visitors describe and emotionally evaluate the Historic Centre of Macao in their multilingual online reviews, in terms of recurring themes, values, and everyday experiences across heritage components, platforms, and languages (RQ1). The results show that reviews are, overall, strongly positive across the system: mean sentiment scores on a 1–5 scale cluster around four stars for most platform–language combinations, with component-level means ranging from 3.27 (Moorish Barracks) to 4.26 (St. Lawrence’s Church).² Even the anchor case of the Ruins of St Paul’s, which attracts the highest volume of reviews, sits close to the system-wide middle (3.95) rather than at the top of the distribution. This broadly positive evaluative pattern is consistent with previous work showing that online reviews of heritage and cultural tourism sites tend to display a positivity bias, with negative experiences less likely to be reported (Huertas and Marine-Roig, 2016; Santa-Cruz and López-Guzmán, 2017). At the same time, the keyterm out-

² See the system-wide sentiment map and component heatmap in Figure 5.13.

puts indicate that these positive sentiments are grounded in concrete aspects of the visit rather than only in generic praise. Across the system, visitors repeatedly refer to architectural and urban features (e.g., churches, squares, gardens, façades), to practical aspects of the visit (admission, entrance, queuing, crowding), and to activity-related terms such as walking, taking photos, praying, or attending festivals.³ In this sense, public reviews portray the Historic Centre as a lived urban landscape that combines visual enjoyment, movement, and social practices, echoing the Historic Urban Landscape perspective, which emphasises everyday use and contemporary life as integral parts of urban heritage (Bandarin and Van Oers, 2012).

The second research question focused on the extent to which these public narratives align with the OUV criteria adopted for the Historic Centre of Macao (RQ2). The two-stage OUV pipeline operationalises this question by first filtering OUV-relevant sentences using a generic classifier and then aligning them to Macao-specific claim embeddings. For the anchor case, the Macao-specific alignment yields a criterion profile strongly dominated by Criterion (iv) (*material urban fabric & architectural ensembles*), with Criterion (iii) (*contact zone & cultural fusion*) as a secondary contribution, and only small residual shares for Criterion (ii) (*cultural–intellectual exchange mechanisms*) and Criterion (vi) (*circulation of ideas & socio-political change*).⁴ System-wide, the pattern is similar: when reviewers express OUV-aligned content, they most often talk about the tangible built environment—streetscapes, church ensembles, urban textures, while more abstract historical and intellectual aspects related to cultural exchange and circulation of ideas are rarely made explicit.⁵ This uneven alignment resonates with critiques of the Authorized Heritage Discourse (Smith, 2006; Meskell, 2018), which argue that expert formulations of heritage values often prioritise certain narratives and scales that are not necessarily shared or articulated in everyday language. In the Macao case, the OUV alignment results suggest that visitors intuitively connect to the material and spatial dimensions of the historic centre and, to a lesser extent, to its role as a contact zone, while the more abstract intellectual and geopolitical dimensions remain largely an expert concern.

The third research question asked how the heritage components of the Historic Centre are visually represented in user-generated images, and what system-wide visual themes and language-sensitive preferences emerge from these photographs (RQ3). The visual clustering analysis identifies eight recurring themes, including *Temples & Folk Shrines*, *Buildings & Streetscapes*, *Lanes, Alleys & Walls*, *Interiors & Deco-*

³ See the system-wide word cloud in Figure 5.14.

⁴ See the anchor case profile 5.2.3.

⁵ See the system-wide criterion distribution and component-level criterion profiles in 5.3.3.

rative Details, Guia Fortress & Lighthouse, Green Spaces & Parks, Catholic Church Interiors, and Facades. At the system level, *Interior & Decorative Details* and *Green Spaces & Parks* are the most frequent themes, followed by *Facades* and *Buildings & Streetscapes*, indicating that visitors often photograph close-range material details and greenery as well as more conventional building views.⁶ Component-level profiles show that some sites are visually dominated by a single motif—such as Ruins of St Paul’s by *Buildings & Streetscapes* and *Facades*, or Guia Fortress by *Guia Fortress & Lighthouse*—while others display more mixed portfolios that combine several themes. Language-normalised theme distributions suggest that, when aggregated across all components, English, Simplified Chinese, and Traditional Chinese images share very similar visual patterns, with only modest shifts in emphasis (for instance, a slightly higher share of interior imagery in the Traditional Chinese subset).⁷ These findings support earlier observations that tourist photography tends to take around a limited set of iconic views and compositional routines (Urry, 1992; Gartner, 1994), while also highlighting the importance of interiors, details, and green spaces in visitors’ visual engagement with the historic centre.

The fourth research question examined how textual and visual UGC complement or diverge from each other in representing the heritage values of the Historic Centre at the component level (RQ4). The cross-modal analysis uses component profiles to relate visual theme dominance to OUV-aligned claim shares and average sentiment. Several non-random associations emerge: components with higher proportions of *Buildings & Streetscapes* imagery tend to have higher shares of Criterion (iv) (*material urban fabric & architectural ensembles*); components dominated by *Catholic Church Interiors* and *Interiors & Decorative Details* are more strongly associated with Criterion (vi) (*circulation of ideas & socio-political change*) and Criterion (iii) (*contact zone & cultural fusion*); and *Green Spaces & Parks* shows a negative association with Criterion (iv) (*material urban fabric & architectural ensembles*).⁸ In terms of evaluative tone, components visually dominated by landmark and streetscape themes are more positively framed than those dominated by transitional elements such as lanes or by greenery alone. At the level of individual components, highly iconic sites such as Ruins of St Paul’s exhibit a strong convergence between what is most often photographed and what is most frequently discussed in OUV-aligned sentences, whereas other components display more complex or weaker cross-modal relationships. Taken together, these patterns indicate that textual and visual representations are neither fully redundant nor independent.

⁶ See the system-wide visual theme distribution and exemplars in 5.4.

⁷ See the language-sensitive visual preferences and component × visual theme profiles in 5.27.

⁸ See the correlation matrices between visual themes, OUV claims, and sentiment in 5.5.

Instead, they capture partly overlapping dimensions of public perception, which aligns with recent arguments in multi-modal heritage research that different media encode different facets of how people experience and value place (Bai et al., 2023).

In the following sections, these answers to the research questions are unpacked in more detail.

6.2. Public Perception and OUV Alignment

A central aim of the thesis was to examine how far public narratives in user-generated reviews correspond to the Outstanding Universal Value (OUV) of the Historic Centre of Macao. OUV is an expert framework that condenses historically complex processes into a small set of criteria intended to be comparable across sites, which means it does not automatically translate into everyday visitor language (Jokilehto, 2008; Smith, 2006; Labadi, 2013). The Macao property is inscribed under cultural criteria (ii), (iii), (iv), and (vi) (Cultural Institute of the Macao SAR Government, 2005), and the Ruins of St Paul’s were used as an anchor case to explore this relationship before moving to system-wide patterns.

Within the Historic Centre, the Ruins of St Paul’s are both historically significant and highly visible in contemporary tourist practice. As the remaining facade and staircase of the former Church of St Paul and St Paul’s College, the site is closely tied to the Jesuit presence in Macao and to longer histories of religious, educational, and cultural exchange in the region (Pinheiro et al., 2005). At the same time, it is one of the most iconic and frequently photographed nodes in the heritage core, and it is repeatedly foregrounded in official heritage and visitor-facing narratives of the Historic Centre (Cultural Affairs Bureau of the Macao SAR Government, 2026; Cultural Institute of the Macao SAR Government, 2005). In the dataset assembled for this thesis, it also attracts the largest volume of reviews and images, making it a useful anchor for interpreting how OUV-related claims enter (or fail to enter) short-form review language.

The Macao-specific OUV alignment for Ruins of St Paul’s shows a clear concentration on Criterion (iv) (*material urban fabric & architectural ensembles*), with Criterion (iii) (*contact zone & cultural fusion*) as a secondary contribution and only small residual shares for Criterion (ii) (*cultural–intellectual exchange mechanisms*) and Criterion (vi) (*circulation of ideas & socio-political change*). This pattern is stable across languages and platforms: when the aligned sentences are broken down by language or by platform, Criterion (iv) remains the dominant share in all subsets,

while (ii) and (vi) stay consistently small. In other words, when visitors write about Ruins of St Paul’s in ways that match the Macao OUV claims, they most often do so by referring to its built form and, to a lesser extent, to ideas of cultural mixture, rather than to its historical role in intellectual networks or wider socio-political change.

Comparing the two stages of the alignment pipeline makes this even clearer. For Ruins of St Paul’s, the generic Stage 1 classifier initially concentrates most OUV-related sentences in Criterion (ii), whereas the Macao-specific Stage 2 alignment strongly re-weights the profile towards Criteria (iv) and (iii). At the sentence level, the majority of Stage 1 Criterion (ii), (iii), (iv), and (vi) sentences are reassigned to Criterion (iv) in Stage 2, with smaller portions moving into Criterion (iii) and only minor shares remaining in (ii) or (vi). This indicates that generic heritage language in visitor reviews is easily interpreted as “cultural exchange” in a broad sense, but once the local claim set for Macao is taken into account, the most specific match for Ruins of St Paul’s lies in its material urban fabric and contact-zone character. Historically, this is consistent with the way the site is described in the nomination dossier and later commentary, where the facade is highlighted as a key example of East–West architectural fusion, while the college’s intellectual role is discussed at a more abstract scale for the city as a whole (Cultural Institute of the Macao SAR Government, 2005; Pinheiro et al., 2005).

At the system-wide level, the same basic imbalance appears. The Stage 2 Macao-specific OUV profile aggregated across all components is again dominated by Criterion (iv), with Criterion (iii) as the second largest share and Criteria (ii) and (vi) remaining comparatively small. When generic Stage 1 outputs are compared to Stage 2, the transition table shows a similar re-weighting as in the anchor case: around 60–90% of sentences initially classified as each generic criterion move to Criterion (iv) in the Macao-specific alignment, with a secondary flow into Criterion (iii) and only limited persistence in (ii) and (vi). Taken together, the anchor-case and system-wide profiles point to a structured but uneven alignment between public narratives and OUV. Visitors do notice values central to Criteria (iv) and (iii): they talk about facades, streetscapes, ensembles, and, in some cases, cultural mixture in the historic core. These dimensions are closely tied to what is seen and experienced on-site and thus more readily enter everyday review language. By contrast, the intellectual, religious, and geopolitical dimensions that underpin criteria (ii) and (vi) are much less visible in the aligned sentences, even though they are prominent in expert OUV formulations for the property (Cultural Institute of the Macao SAR Government, 2005; Labadi, 2013). This gap does not imply that visitors reject these values; rather, it highlights that they are difficult to express within the short,

experience-focused format of online reviews and within platform conventions that privilege immediately observable cues (Smith, 2006).

At the same time, public narratives surface themes that are largely absent from the OUV framework, such as crowding, commercialisation, and the presence of snack streets around Ruins of St Paul’s and other high-traffic components. These concerns reflect contemporary pressures on the Historic Centre as a tourism destination and as part of Macao’s everyday urban life, and they influence how the site is perceived and evaluated. For heritage management, this implies a dual challenge. On the one hand, there is a need to communicate the less visible OUV dimensions in ways that connect to visitors’ existing experiences of the urban fabric (for example, by linking architectural ensembles to narratives about exchange, education, or religious practice that are locally grounded). On the other hand, managers must engage with present-day issues raised in user-generated content that lie outside the original inscription rationale but nonetheless shape public perception of the property. In this sense, the OUV alignment results should be read not only as an assessment of correspondence, but also as a diagnostic of where interpretive “translation” succeeds (material fabric and legible fusion cues) and where it remains weak (abstract intellectual and geopolitical dimensions) under current platform conditions (Jokilehto, 2008; Labadi, 2013).

6.3. Language and Platform Differences

The results show that both language and platform systematically shape how visitors describe and evaluate the Historic Centre of Macao. This is not simply a sampling issue; it reflects a broader point made in tourism and platform research: user-generated heritage narratives are produced within platform-specific incentive structures, interfaces, and audiences, which in turn influence what becomes sayable, salient, and comparable (Van Dijck et al., 2018; Xiang and Gretzel, 2010). Treating all reviews as interchangeable “public opinion” therefore risks conflating site perception with the socio-technical conditions of review production.

In this thesis, platform differences appear consistently across sentiment distributions, criterion composition, and keyword rank patterns. A plausible interpretation is that Google Maps and TripAdvisor occupy different positions in the travel information ecosystem. Prior work suggests that Google Reviews are often produced in close temporal proximity to a visit and under mobile, place-based conditions, whereas TripAdvisor reviews more often reflect deliberate, post-trip narration aimed at an imagined community of future travellers (Xiang and Gretzel, 2010; Mellinas and

Sicilia, 2024). These differences in writing situation can shift what is emphasized: quick evaluative signals versus more structured accounts that connect attractions to itineraries, expectations, and comparative benchmarks.

For Chinese-language platforms, the literature similarly cautions that domestic review ecosystems are not merely linguistic variants of Western platforms. Reviews on Chinese travel platforms can foreground different practical concerns and destination-image elements, and cross-platform comparisons within the Chinese ecosystem show meaningful divergences in what attributes are discussed and how destinations are framed (Guo et al., 2021; Leung et al., 2011). Against this backdrop, the platform gaps observed here should be understood as structured differences in communicative purpose (why people write), not only as differences in tourist composition.

Language groups also differ in the distribution of sentiments and in the relative emphasis placed on OUV-aligned criteria. Part of this likely reflects audience composition. Official statistics indicate that Macao tourism is dominated by arrivals from the Greater China region, with Mainland visitors forming the largest share, while international arrivals represent a smaller proportion (Dioko et al., 2021; Government of the Macao Special Administrative Region, 2026). As a result, English-language reviews (more likely to be produced by international visitors and internationally oriented residents) and Chinese-language reviews (more likely to reflect Greater China mobility) may encode different baseline expectations of urban heritage, crowding, religious symbolism, and “must-see” itineraries.

Macao’s heritage context makes these differences especially consequential. The Historic Centre is commonly interpreted through the lens of long-term Sino-Portuguese interaction and Catholic material heritage, but it is also embedded in a contemporary tourism economy dominated by short-stay, high-density visitation and a regionally integrated mobility regime. In practice, this means different visitor publics may mobilize different interpretive scripts: some lean toward historical “fusion” narratives and landmark symbolism, while others foreground route efficiency, density management, and practical visit utility. The language effects in keyword ranking divergence and their statistical significance (Figure 5.17) are consistent with this idea: linguistic groupings do not merely translate the same experience; they can re-weight which aspects of place are reportable and meaningful.

A key takeaway is that language- and platform-conditioned differences are not noise to be averaged out; they are part of the phenomenon. This aligns with broader warnings in tourism analytics that single-platform conclusions may be biased because platforms act as distinct “information agents” that selectively represent destination attributes (Guo et al., 2021). For heritage management, the practical implication is

that “public perception” should be treated as plural and segmented: what is legible on Google Maps may not be legible on TripAdvisor, and what is emphasized in English may not match what is emphasized in Simplified or Traditional Chinese.

Finally, these results also speak back to the OUV alignment analysis. If different platform publics foreground different evaluative logics, then observed variation in criterion composition should be interpreted as a mapping between (i) the site’s value propositions and (ii) platform-specific publics and their communicative norms (Van Dijck et al., 2018; Zeng and Gerritsen, 2014). This reinforces the rationale for a cross-platform, multilingual design: it helps distinguish stable, site-wide value signals from platform-conditioned emphases, and it clarifies which heritage narratives are amplified or mute within particular digital infrastructures.

6.4. Cross-modal interpretation: convergence and divergence

Bringing the textual and visual analyses together shows that representations of the Historic Centre of Macao are cross-modally organised rather than fragmented. At the component level, visual theme proportions, OUV-aligned claim shares, keyterm signals, and average sentiment co-vary in structured ways. In particular, the Spearman correlation patterns indicate that certain visual motifs are repeatedly associated with specific OUV criteria (see Figure 5.29, and that these correspondences are mirrored lexically in the words visitors use to label and explain what they see.

The strongest cross-modal convergence emerges around highly photographable and culturally legible heritage elements. Components dominated by *Buildings & Streetscapes* and *Facades* tend to exhibit higher shares of Criterion (iv) (*material urban fabric & architectural ensembles*) and slightly higher average sentiment in the component-level comparisons. Lexically, these exterior-dominant components co-vary with built-heritage descriptors (e.g., *Colonial*, *Archway*), indicating that canonical photographic framings are echoed in the vocabulary used to describe the same components.

This convergence aligns with long-standing work on the tourist gaze, which argues that tourists are socialised to seek out shareable, frameable scenes—especially facades, vistas, and recognisable “postcard” compositions—that can be easily captured and circulated (Urry, 1992; Gartner, 1994). In Macao, this gaze is reinforced by official heritage and interpretation materials that repeatedly foreground a small set of iconic compositions (notably the stepped approach and facade of the Ruins of St Paul’s,

the colonnades and paving patterns of Senado Square, and prominent church fronts), making it unsurprising that these components not only attract dense photographic attention but also generate text that aligns most clearly with Criterion (iv).

A different but equally interpretable alignment appears for interior-focused imagery. Components where *Catholic Church Interiors* and *Interiors & Decorative Details* dominate are more strongly associated with Criterion (vi) (*circulation of ideas & socio-political change*) and Criterion (iii) (*contact zone & cultural fusion*). Because Criterion (vi) is about ideas and historical influence, visitors often cannot name it directly; however, altars, iconography, plaques, and museum-style displays provide concrete visual cues that may help such meanings appear in the alignment results. Visitors can take pictures of altars, sculptures, ceiling paintings, tiles, and other decorative components in these sections. These aspects draw attention to symbolic atmospheres, religious practices, and culturally hybrid materials.

The lexical layer is consistent with this separation: Catholic interior and facade themes co-vary with *Church* and *Portugal*, whereas temple-related imagery co-varies with *Temple*, *China*, and *Incense*, indicating a stable religious-cultural axis visible across modalities. In this sense, interior imagery can be read as a visual counterpart to textual narratives that focus on the symbolic and ceremonial functions of the built environment, even when deeper historical intellectual networks remain implicit in short review formats.

Cross-modal divergence is clearest for components visually dominated by *Green Spaces & Parks* or *Lanes, Alleys & Walls*. These themes are weakly negatively associated with sentiment and show weaker correspondence with Criterion (iv), suggesting that when such transitional elements dominate the visual framing, reviews shift toward more mixed and practical concerns (e.g., shade, rest, access difficulty, maintenance) rather than toward explicit recognition of architectural ensembles.

From a Historic Urban Landscape (HUL) perspective, these spaces matter: paths, small gardens, and connective lanes contribute to everyday life and continuity of urban form, even if they are not central in official “heritage sight” imaginaries (Bandarín and Van Oers, 2012). The observed divergence therefore does not imply that such spaces lack value; rather, it indicates a mismatch between (i) what visitors notice and photograph as part of moving through the historic centre and (ii) what they frame as OUV-relevant when writing reviews. In other words, some of the most practically salient aspects of visiting Macao’s heritage landscape appear to remain primarily textual (and often experience-oriented) rather than entering the formal OUV alignment channel.

The Ruins of St Paul’s illustrates how these system-wide tendencies concentrate at a single, iconic component. Its image set is dominated by exterior themes (*Buildings & Streetscapes* and *Facades*), and its Macao-specific OUV alignment is strongly concentrated on Criterion (iv), with Criterion (iii) as a secondary share, alongside consistently positive sentiment across languages and platforms. Its keyterm profile similarly reflects this iconic framing, tracking the dominant exterior motifs rather than more diffuse experiential cues. This pattern is consistent with the site’s long-established symbolic status in Macao’s heritage narratives and with the way visitors reproduce that status through both photographs and short comments.

Taken together, the correlations suggest a division of labour between images and text. Images privilege what is immediately photographable and socially shareable (facades, vistas, iconic objects), while text more readily accommodates temporal sequences, constraints, and evaluative explanations. This helps explain why the strongest cross-modal correspondences cluster around concrete landmarks and culturally legible motifs, whereas more processual concerns (crowding, access, comfort) remain primarily textual even when they shape the experience of the same components. This complementarity resonates with multimodal heritage analytics arguing that different media encode different facets of how people experience and value place (Bai et al., 2023).

For interpretation, the cross-modal synthesis suggests that the Historic Centre is represented through shared routines that tie together “what is photographed” and “what is said”, especially for iconic streetscapes and facade-dominant sites. At the same time, divergence highlights where visitor experience is organised around everyday urban infrastructure and comfort rather than around formal value language. For heritage management, this implies that efforts to communicate less visible OUV dimensions, or to redistribute visitor attention beyond a small set of postcard compositions, cannot ignore the visual habits embedded in tourist photography and platform circulation.

6.5. Methodological Reflections and Limitations

The findings in this thesis rely on a multi-step analytical pipeline that connects user-generated reviews and images to UNESCO’s Outstanding Universal Value (OUV) framework. While this design makes it possible to work across languages, platforms, and modalities, it also introduces several methodological constraints. This section reflects on the main choices made in the study and discusses their implications for interpreting the results.

6.5.1. Data sources, Sampling, and Representativeness

The analysis is based on user-generated content from Google Maps, TripAdvisor, and Ctrip, restricted to the officially designated heritage components in the Historic Centre of Macao. For each component, up to the top 100 most relevant reviews per available language were collected from Google Maps, together with English-only reviews from TripAdvisor and simplified Chinese reviews from Ctrip, and images were limited to photos directly attached to Google Maps reviews. This strategy ensures a coherent, place-anchored corpus that can be compared across platforms and languages, but it also creates systematic biases in coverage. Highly visible components such as the Ruins of St Paul’s attract dense review and image activity, whereas less well-known sites are represented by only a small number of entries. The data summary in Chapter 3 and the overview in Chapter 5 already show that review and image volumes vary substantially across components, with some sites having around 70 photographs and others more than 700. Components with limited material are therefore treated as lower-evidence cases and their profiles should be interpreted with greater caution.

Using only three platforms also shapes the picture of “public perception” that emerges. Google Maps provides a multilingual baseline, but its ranking of “Most relevant” reviews is influenced by opaque platform algorithms. TripAdvisor mainly captures independent travellers writing in English, often with more elaborate reflections on heritage quality and interpretation. Ctrip focuses on mainland Chinese tourists and combines practical travel advice with experiential comments. These groups overlap but are not identical, and previous work has shown that platform audiences differ in socio-demographic composition and travel behaviour. As a result, the study does not capture all visitor segments equally. For example, organised tour groups who do not leave online reviews, or visitors using other social media platforms, are largely absent.

Language coverage is another source of partial representation. The corpus focuses on English, simplified Chinese (zh-CN), and traditional Chinese (zh-TW), which together cover major visitor groups in Macao, but reviews written in Portuguese or other languages are not included. Moreover, language cannot be treated as a simple indicator for nationality or culture: English reviews can be written by both native and non-native speakers, and Chinese-language reviews come from a wide range of backgrounds. The analysis therefore treats languages as communicative choices rather than as direct indicators of origin, which limits how far the results can be generalised to specific national groups.

Therefore, these constraints mean that the results are best read as comparative patterns within a particular slice of online engagement, rather than as a complete description of all perceptions of the Historic Centre. The thesis already stresses this point in the Results chapter, where component-level plots are framed as comparative indicators rather than exhaustive statements about what each site is.

Finally, the findings should be interpreted as patterns in digitally mediated public expression rather than as an exhaustive account of stakeholder values. Interviews, surveys, or participatory workshops with residents, local practitioners, and repeat visitors would likely surface additional perspectives that are weakly represented in platform reviews, including more ambivalent or contested meanings, longer-term attachments, and governance-related concerns. Such methods would not necessarily overturn the cross-modal regularities observed here, but they could broaden the value spectrum and help distinguish platform-driven conventions from more durable place meanings.

6.5.2. Textual Pipeline: Sentiment and Experiential Signals

On the textual side, sentiment analysis and keyterm extraction were used to summarise how visitors feel about the heritage components and what they emphasise in their written descriptions. Sentiment analysis is based on a multilingual BERT classifier built on XLM-RoBERTa and fine-tuned on generic review data. This choice is appropriate for a multilingual environment like Macao, where reviews contain English, simplified Chinese, and traditional Chinese, sometimes within the same text, and existing work has shown that such models perform well on short, informal comments. At the same time, the model was not specifically trained on Macao heritage reviews, so subtle emotional cues, local idioms, or culturally specific expressions may be misclassified, and the positive–neutral–negative framing cannot fully capture mixed or ambivalent feelings. The decision to operate at review level simplifies the emotional signal, but it also means that within-review contrasts (for example, positive comments about architecture combined with negative remarks about crowding) are averaged into a single score.

Keyterm extraction relies on KeyBERT with a multilingual SentenceTransformer model (paraphrase-multilingual-mpnet-base-v2). This configuration is well suited for short, multilingual reviews and helps to identify semantically central phrases rather than just frequent words. However, several design choices introduce limitations. First, a dedicated stopwords list was used to remove all occurrences of component names in English and Chinese from the candidate keyphrases. This

avoids outputs dominated by place names, but it also suppresses expressions where the name itself carries experiential meaning, such as “take photos at the Ruins of St Paul’s.” This potential information loss is acknowledged in the Methodology chapter and remains a structural limitation of the keyterm results. Second, automatic keyterm ranking cannot distinguish between evaluative and descriptive uses of the same term. Phrases like “crowded streets” may signal either a lively atmosphere or an uncomfortable experience, depending on context, which is not fully visible in the aggregated outputs.

Pair-wise comparisons of keyterm rankings across languages and platforms use the non-parametric Mann–Whitney U test. This approach is robust to non-normal distributions and allows for rank-based comparisons without assuming equal variances. At the same time, only shared keyterms enter these tests, and the number of overlapping phrases can be limited for smaller subsets. As a result, some of the statistical contrasts in the heatmaps should be interpreted as indicative tendencies.

6.5.3. OUV Alignment: Modelling Heritage Narratives

The two-stage OUV alignment pipeline is a central part of the thesis. Stage 1 applies the WHOSe Heritage classifier, which combines BERT and ULMFiT models trained on UNESCO documentation and heritage-related texts by Bai et al. (2021), to label review sentences with generic OUV criteria. Stage 2 then aligns these sentences to a Macao-specific OUV knowledge base using Sentence-BERT embeddings and a cosine similarity threshold. This design has clear strengths: it builds directly on current heritage NLP methods and allows a large, multilingual review corpus to be analysed against formally defined OUV claims. The use of sampling-based validation in both stages and the explicit tuning of thresholds (for example, the similarity threshold of $T = 0.259$ in Stage 2) provide a degree of transparency and control over precision and recall trade-offs.

At the same time, several limitations still exist. First, the Stage 1 classifier is trained on global UNESCO documentation, not on Macao alone. Although the validation results on the 90-sentence sample are encouraging, some generic OUV-related language in the reviews may still be mislabelled, especially in more ambiguous or narrative paragraph. Second, Stage 2 relies on machine-translated English for all Chinese sentences, and on a sentence embedding model trained on general purpose text. Even though additional segmentation was applied after translation to keep sentence boundaries aligned, there is always a risk that translation shifts meanings or softens culturally specific references, particularly for terms that describe local

practices, religious concepts, or food.

Third, the thresholding strategy for Stage 2 favours high confidence matches and removes low similarity pairs that showed poor accuracy in manual checks. This improves the reliability of the aligned dataset but also reduces volume and may underrepresent more implicit or creative ways in which visitors talk about OUV-related aspects of the site. The Macao specific alignment also concentrates much of the signal in Criterion (iv) (*material urban fabric & architectural ensembles*), with Criterion (iii) (*contact zone & cultural fusion*) as a secondary share, while Criteria (ii) and (vi) are less frequent. Part of this pattern likely reflects genuine emphasis on tangible built heritage in visitor narratives, but part of it may be driven by how the OUV knowledge base is formulated and how the embedding model measures similarity. The study follows the official decision to focus on criteria (ii), (iii), (iv), and (vi), leaving out earlier discussions of criterion (v) in the nomination dossier. This means that potential OUV-related content linked to landscape or ecological aspects is not explicitly modelled.

Finally, the validation exercises in both stages are based on relatively small samples of 90 sentences each. They provide useful checks on decision thresholds but do not cover every possible error pattern, and the OUV alignment should be interpreted as an explainable, structured estimate of how reviews relate to the official criteria—not as definitive ground truth.

6.5.4. Visual Themes and Cross-modal Alignment

On the visual side, the study uses a pretrained vision encoder to embed 14,079 images and applies K-means clustering to derive eight global visual themes. These themes, such as *Buildings & Streetscapes*, *Facades*, *Green Spaces & Parks*, and *Catholic Church Interiors*, prove interpretable at both system and component level, and they align well with motifs identified in previous work on tourist gaze and visual destination images. However, the approach also inherits limitations from generic computer vision practice. The encoder is trained on large, general-purpose image datasets rather than on heritage-specific corpora, and the K-means algorithm assumes spherical clusters in the embedding space. The t-SNE plots in the Results chapter already show that outdoor built-environment themes overlap in feature space, and misassignments at the margins between themes are likely.

The cross-modal comparison relies on component-level Spearman correlations between visual theme shares and OUV-claim shares, as well as between visual dominance and sentiment. This design captures rank-based structural alignments but

does not account for all sources of dependence in the data. For example, components with very similar visual profiles but different visitor volumes contribute equally to the correlation, and the analysis does not explicitly model spatial autocorrelation or network effects between nearby components. The observed associations should therefore be understood as indications of systematic co-occurrence patterns rather than as causal relationships between what is photographed and what is said.

6.5.5. Influences for Interpretation

Despite these limitations, the methodological design provides a transparent and reproducible way to connect public online narratives with UNESCO OUV frameworks at multiple scales. The combination of multilingual sentiment analysis, semantic keyterm extraction, a validated two-stage OUV pipeline, and global visual themes allows the thesis to move beyond isolated case studies and to examine systematic patterns across the heritage system. At the same time, the reliance on specific platform user-generated content, pretrained NLP and vision models, and threshold-based alignment means that the results are always affected by technical choices.

For interpretation, this implies that the thesis should be read as an analysis of how a specific group of digitally active visitors currently describe and depict the Historic Centre of Macao on today's major platforms. Building on this starting point, future work could broaden the evidence base by adding more data sources, involving local stakeholders directly in annotation and validation, and testing alternative modelling and representation choices. With these extensions, the framework developed here can evolve into a more participatory and context-aware toolkit for studying public perceptions of World Heritage in contemporary digital environments.

6.6. Implications for World Heritage Management and Future Work

The findings of this thesis have some practical implications for how the Historic Centre of Macao is managed and communicated as a World Heritage site. At the most basic level, the review and image data confirm that visitors clearly recognise core elements of the site's Outstanding Universal Value (OUV), especially the visual mixture of Portuguese and Chinese architecture and the dense ensemble of churches, temples and urban squares in the old town. This broadly matches the official OUV statement, which frames Macao as a living testimony of sustained East–West cultural

exchange in an urban setting (UNESCO World Heritage Centre, 2005). In that sense, the user-generated content analysed here supports the way Macao is presented in UNESCO and local government documents, and offers independent evidence that these messages are reaching international visitors.

At the same time, the OUV alignment analysis also suggests where there are gaps. Criteria related to everyday social life, local traditions or religious meanings appear less often in reviews than more visual aspects such as iconic façades, night views or casino skylines. This echoes wider debates in critical heritage studies which argue that official narratives tend to privilege monumental architecture and expert voices, while more ordinary or community-based values remain under-represented (Smith, 2006). For Macao, this means that heritage managers may need to invest more in communicating the social and spiritual dimensions of the Historic Centre, such as local Catholic and Chinese festivals, neighbourhood networks around the squares, or memories of older residents, so that these aspects become more visible in visitor accounts as well.

The digital traces used in this thesis can also complement existing monitoring and management tools. The Operational Guidelines stress that States Parties must not only protect OUV but also put in place ongoing mechanisms to monitor changes, manage tourism and involve local stakeholders (UNESCO World Heritage Centre, 2024). Macao has already developed a Protection and Management Plan for the Historic Centre under Article 51 of the Cultural Heritage Protection Law and uses impact assessments for projects that might affect OUV (Macao Special Administrative Region, 2013). Periodic analysis of reviews and images, using approaches similar to those in this thesis, could become a low-cost supplement to such monitoring. For example, sudden shifts in sentiment for specific squares, or recurring complaints about crowding, noise, or maintenance, could serve as early warning signals for the Cultural Affairs Bureau and tourism authorities.

The results also speak to current efforts to apply the Historic Urban Landscape (HUL) approach in World Heritage cities. The HUL Recommendation encourages cities to treat heritage as part of a wider urban system and to engage local communities in identifying key values and negotiating future change. (UNESCO, 2011) In Macao, reviews and images show how residents and visitors use the Historic Centre not only as a heritage site but also as a place for shopping, street food, religious practice and everyday mobility. These patterns can inform urban planning decisions, for example, how to balance pedestrian flows, local commerce and residential needs in streets around Senado Square or the Ruins of St Paul's, so that the site remains both liveable for residents and meaningful for visitors.

From a tourism perspective, the analysis of platform and language differences highlights the need for more targeted communication strategies. International research based on TripAdvisor and other platforms shows that although UNESCO label encourages tourists' interests, but does not automatically translate into more sustainable behaviour (Koufodontis and Gaki, 2022; Van der Zee et al., 2024). In the Macao data, some user groups focus strongly on "Instagrammable" views and shopping, while others pay more attention to history or spirituality. Management agencies could therefore design platform-specific messages, for instance, short, visually oriented explanations on Google Maps and Ctrip, or more in-depth narratives and links on TripAdvisor and official websites, to reach different audiences in their own languages. This would also respond to the Cultural Heritage Protection Law's emphasis on education and public awareness as part of safeguarding duties (Macao Special Administrative Region, 2013).

Finally, this thesis shows that digital user-generated content can be used as a bridge between local heritage policies and global debates on sustainable tourism and community participation. Recent UNESCO and academic work calls for World Heritage cities to move away from a purely growth-oriented tourism model and to involve residents more actively in decisions about visitor numbers, public space and housing. (UNESCO, 2011; Rasoolimanesh et al., 2017) Similar tensions also exist in Macao, where the Historic Centre is both a tourism icon and a lived neighbourhood (Lai and Hitchcock, 2017). By systematically analysing what different groups say about the site online, heritage managers gain another channel to hear concerns, identify conflicts and test new communication approaches. In that sense, the methods developed here could be adapted not only for future monitoring in Macao but also for other urban World Heritage sites that face similar challenges of mass tourism, rapid urban change and unequal visibility of local voices.

7. Conclusion

7.1. Summary of Aims and Approach

This thesis set out to analyse how the Historic Centre of Macao is described, evaluated, and visually represented in user-generated content on major travel and map platforms. In particular, the thesis treats platform environments as part of the production of public heritage meanings, because destination images are continuously reproduced through platform-specific conventions and multimodal narratives rather than only through official interpretation. The main objective was to understand how visitors describe their experiences of the site in everyday language and images, and how these public expressions relate to the Outstanding Universal Value (OUV) criteria that underpin Macao’s World Heritage status. By bringing together multilingual textual reviews, user-uploaded images, and UNESCO documentation, the study aimed to bridge the gap between expert narratives about heritage and the ways in which visitors actually talk about and see the Historic Centre.

To achieve this, the thesis developed a place-based analytical framework that operates at the level of the officially designated heritage components. On the textual side, reviews from Google Maps, TripAdvisor, and Ctrip were pre-processed, segmented into sentences, and analysed with a multilingual sentiment model and a keyterm extraction pipeline. A two-stage OUV classifier was then used to filter OUV-relevant sentences and align them with a small set of Macao-specific criteria derived from official documents. On the visual side, a pretrained vision encoder and K-means clustering were used to derive eight interpretable visual themes from 14,079 user-generated images. Component-level profiles of sentiment, keyterms, OUV claims, and visual themes were finally linked to explore structural relationships between what visitors photograph and what they write.

The analysis followed a two-step logic. Ruins of St Paul’s, as the most visible and symbolically important component in the Historic Centre, was used as an anchor case to make the pipeline transparent and to calibrate thresholds and labels. The same procedures were then applied to the full system of 27 components to derive

system-wide patterns, compare language and platform subsets, and examine cross-modal alignments between text and images.

7.2. Synthesis of Main Findings

Across the full corpus, the textual analysis shows that visitors' written perceptions of the Historic Centre of Macao are strongly positive. Mean sentiment scores cluster around four stars on a five-point scale, with only modest variation between components and between platform and language combinations. Even for Ruins of St Paul's, which receives the largest volume of reviews and is often crowded, the average sentiment sits near the system-wide middle rather than at the bottom of the distribution. This confirms that, in everyday discourse, the Historic Centre is largely framed as a pleasant and worthwhile place to visit.

Keyterm patterns provide a more detailed picture of what underlies this positivity. System-wide, visitors repeatedly refer to churches, squares, gardens, streets, and facades, but also to practical aspects of the visit such as queues, admission, views, and photo-taking. In this sense, the Historic Centre appears in reviews as a lived urban landscape where architectural ensembles, movement, and social practices are closely interconnected. Language- and platform-specific profiles suggest that "public perception" is not a single signal but a segmented one: different platform publics and writing situations foreground different aspects of the visit, even when they converge on a shared core of Macao's Sino-Portuguese heritage identity.

The OUV alignment results reveal a structured but uneven correspondence between public narratives and the official criteria adopted for the Historic Centre. For both the anchor case and the system-wide profiles, Criterion (iv), which emphasises material urban fabric and architectural ensembles, clearly dominates the aligned signal, with Criterion (iii) on cultural fusion as a secondary contribution. Criteria (ii) and (vi), which relate to broader cultural-intellectual exchange and socio-political change, are only weakly expressed in visitor sentences. Where Criterion (vi) does emerge, it appears largely through indirect, experience-near cues (e.g., religious interiors and symbolic objects) rather than through explicit references to broader historical processes. This pattern suggests that visitors do notice and describe the built form and the visual mixture of Portuguese and Chinese elements that are central to Macao's values. At the same time, more abstract or historically complex values remain difficult to articulate within the short, experience-focused format of online reviews.

Importantly, public narratives also surface themes that are largely absent from the OUV framework, such as crowding, commercialisation, street food, and the everyday use of squares and lanes. These concerns reflect the pressures of tourism and the fact that the Historic Centre functions as a heritage site and as part of Macao's contemporary urban life at the same time. They influence how visitors evaluate the site even though they are not part of the original nomination criteria.

The visual analysis complements these textual findings. Across the system, a small number of visual themes, particularly *Interiors & Decorative Details*, *Green Spaces & Parks*, *Facades*, and *Buildings & Streetscapes*, account for most visitor imagery. The distribution of themes is not random: different components attract different dominant visual emphases, and language subsets show only modest shifts rather than completely different patterns. This indicates a relatively stable tourist gaze and a set of repeatable visual motifs that contribute to destination-image reproduction under contemporary platform circulation.

When textual and visual signals are considered together at the component level, several consistent alignments emerge. Components visually dominated by landmark facades and streetscapes tend to have higher shares of Criterion (iv)-related OUV claims and slightly more positive sentiment. Components where interior-focused themes are more prominent are more strongly associated with Criteria (iii) and (vi), and they sometimes show more mixed evaluative tones. Green spaces and transitional elements such as lanes and walls are more weakly connected to OUV-aligned narratives and often associated with more moderate sentiment. Ruins of St Paul's, Senado Square, key church fronts, and certain temples illustrate these patterns most clearly: they attract both intense photographic attention and rich textual descriptions that foreground the same architectural and urban values.

Overall, the thesis shows that textual and visual user-generated content are neither redundant nor independent. They capture partially overlapping dimensions of public perception, and their structured relationships with OUV criteria. They also suggest that visitors' everyday representations of the Historic Centre do reinforce core elements of Macao's World Heritage profile, while also introducing additional concerns that matter for how the site is experienced.

7.3. Contributions, Limitations, and Outlook

Beyond the specific case of Macao, the thesis makes three main contributions. First, it demonstrates how multilingual reviews and user-generated images from several

platforms can be organised in a coherent, component-based framework and analysed with natural language processing and computer vision tools. The combination of sentiment analysis, keyterm extraction, a validated two-stage OUV alignment pipeline, and global visual themes offers a transparent way to link everyday narratives to formal heritage value categories at scale. Conceptually, it also positions UGC as part of destination-image formation in heritage contexts, showing how everyday multimodal representations relate to and sometimes diverge from formal value frameworks such as OUV.

Second, the study provides empirical evidence on how public perceptions of a World Heritage city relate to its official criteria. It shows that visitors unintentionally reproduce some of the key values emphasised in official documents, especially those related to architectural ensembles and cultural fusion, while other dimensions remain less visible in online discourse. At the same time, it highlights how crowding, commercial pressures, and everyday uses of space shape perceptions in ways that are not captured by OUV statements but are highly relevant for sustainable management.

Third, the work illustrates how digital user-generated content can support heritage practice. The methods used here could, with appropriate safeguards, be integrated into periodic monitoring of sentiment and themes at the component level, help to identify emerging issues, and inform more targeted communication strategies across different languages and platforms. They also speak to wider debates on the Historic Urban Landscape approach by showing how heritage components are embedded in daily mobility, commerce, and social life.

These contributions need to be read alongside several limitations. The study relies on platform-specific, volunteer-contributed data and therefore reflects the views of digitally active visitors rather than the whole population of residents and tourists. The textual and visual analyses are based on pretrained models and threshold choices that introduce their own biases and uncertainties, and the OUV alignment results are best understood as structured approximations rather than as definitive ground truth. In addition, the analysis is based on observable digital traces and cannot fully recover the motivations or meanings that visitors and residents attach to their statements and images. Complementary qualitative methods, such as interviews, focus groups, or short on-site surveys, could help clarify how different publics interpret key themes and why certain values are emphasised or omitted. Finally, the focus on the current configuration of platforms and models means that the findings capture a particular moment in Macao's evolving digital and urban history. Because platform interfaces and ranking systems change over time, the observed language/platform differences should be interpreted as contingent on the platform conditions under which the

traces were produced.

Future work can build on this foundation in several directions. On the data side, additional platforms, languages, and locally curated sources could be added to capture a broader range of perspectives, including those of residents and under-represented groups. Methodologically, adapted language and vision models, as well as more advanced cross-modal techniques, could refine the alignment between user-generated content and heritage value frameworks. From a governance perspective, more participatory approaches, such as involving local stakeholders in annotation and interpretation, or combining online traces with on-site surveys and workshops, would help connect digital evidence with community-based knowledge and decision-making.

Overall, the thesis shows that user-generated reviews and images are a rich resource for understanding how World Heritage sites like the Historic Centre of Macao are experienced, valued, and debated in everyday life. By treating these digital traces as platform-conditioned expressions and organising them within a structured, component-based framework, the thesis offers both conceptual and practical tools for making public heritage meanings more visible in research and management.

References

- Retnasih S. Adiwibowo, Pribadi Widodo, and Imam Santosa. 2015. Correlations between public appreciation of historical building and intention to visit heritage building reused as retail store. *Procedia – Social and Behavioral Sciences*, 184:357–364.
- Dora Agapito, Patrícia Oom do Valle, and Júlio da Costa Mendes. 2013. The cognitive-affective-conative model of destination image: A confirmatory analysis. *Journal of Travel & Tourism Marketing*, 30(5):471–481.
- Asif Agha. 2011. Meet mediatization.
- Azhar Alam, Ririn Tri Ratnasari, Muhammad Hanif Al Hakim, Siti Mutmainah, and Azhar Mufidah. 2024. Comparative netnographic analysis of sharia hotel visitor reviews between tripadvisor and google review. *Multidisciplinary Reviews*, 7(7):2024151–2024151.
- N. Bai, M. Ducci, R. Mirzikashvili, P. Nourian, and A. Pereira Roders. 2023. Mapping urban heritage images with social media data and artificial intelligence: A case study in testaccio, rome. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLVIII-M-2-2023, pages 139–146.
- N. Bai, P. Nourian, R. Luo, and A. Pereira Roders. 2022a. 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):469.
- Nan Bai, Renqian Luo, Pirouz Nourian, and Ana Pereira Roders. 2021. 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*, pages 366–384.
- Shuyang Bai, Bige Tuncer, Mir Abolfazl Mostafavi, and Harvey J. Miller. 2022b. Heri-graphs: A dataset creation framework for multi-modal machine learning on graphs of heritage values and attributes with social media. *Journal on Computing and Cultural Heritage*, 15(4):1–27.
- Seyhmus Baloglu and Ken W McCleary. 1999. A model of destination image formation. *Annals of tourism research*, 26(4):868–897.
- Francesco Bandarin and Ron Van Oers. 2012. *The historic urban landscape: managing heritage in an urban century*. John Wiley & Sons.

- Jan Blommaert, Helen Kelly-Holmes, Pia Lane, Sirpa Leppänen, Máiréad Moriarty, Sari Pietikäinen, and Arja Piirainen-Marsh. 2009. Media, multilingualism and language policing: An introduction. *Language Policy*, 8(3):203–207.
- John D Boy and Justus Uitermark. 2017. Reassembling the city through instagram. *Transactions of the Institute of British Geographers*, 42(4):612–624.
- Bill Bramwell and Liz Rawding. 1996. Tourism marketing images of industrial cities. *Annals of Tourism research*, 23(1):201–221.
- Taina Bucher, Anne Helmond, et al. 2018. The affordances of social media platforms. *The SAGE handbook of social media*, 1(1):233–253.
- Fulvia Calcagni, Júlia Nogué Batallé, Francesc Baró, and Johannes Langemeyer. 2022. A tag is worth a thousand pictures: A framework for an empirically grounded typology of relational values through social media. *Ecosystem Services*, 58:101495.
- Ricardo Campos, Vinícius Mangaravite, Antonio Pasquali, Alípio Mário Jorge, Célia Nunes, and Adam Jatowt. 2020. YAKE! Keyword Extraction from Single Documents Using Multiple Local Features. *Information Sciences*, 509:257–289.
- R. Catelli, L. Bevilacqua, N. Mariniello, V. S. Di Carlo, M. Magaldi, and H. Fujita. 2023. A new italian cultural heritage data set: Detecting fake reviews with bert and electra leveraging the sentiment. *IEEE Access*, 11:52214–52225.
- Irene Cenni and Patrick Goethals. 2017. Negative hotel reviews on tripadvisor: A cross-linguistic analysis. *Discourse, Context & Media*, 16:22–30.
- Irene Cenni and Patrick Goethals. 2020. Positive reviews on tripadvisor: a cross-linguistic study of contemporary digital tourism discourse. *Onomázein*, pages 18–40.
- Yaxin Chen, Chuanchun Hong, Yifan Yang, Jiaxin Li, Yu Wang, Tianyu Zheng, Yinke Zhang, and Feng Shao. 2024. Mining social media data to capture urban park visitors’ perception of cultural ecosystem services and landscape factors. *Forests*, 15(1):213.
- Yi-Fan Chen and Rob Law. 2016. A review of research on electronic word-of-mouth in hospitality and tourism management. *International Journal of Hospitality & Tourism Administration*, 17(4):347–372.
- Jaeyoung Choi, Kaysi Eastlick Kushner, Judy Mill, and Daniel WL Lai. 2012. Understanding the language, the culture, and the experience: Translation in

- cross-cultural research. *International Journal of Qualitative Methods*, 11(5):652–665.
- Alton Yeow-Kuan Chua and Snehasish Banerjee. 2013. Reliability of reviews on the internet: The case of tripadvisor. In *Proceedings of the 5th International Conference on Agents and Artificial Intelligence (ICAART)*, pages 430–435.
- F. Clarizia, F. Colace, M. De Santo, M. Lombardi, and F. Pascale. 2018. A sentiment analysis approach for evaluation of events in field of cultural heritage. In *2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pages 120–127. IEEE.
- Cultural Affairs Bureau of the Macao SAR Government. 2026. Ruins of st. paul's college (former mater dei church, forecourt and staircase). <https://www.wh.mo/en/site/detail/18>. Official site description of the Ruins of St. Paul's, accessed 2026-01-05.
- Cultural Institute of the Macao SAR Government. 2005. *The Historic Centre of Macao*. Cultural Institute of the Macao Special Administrative Region Government, Macao. Includes chapter “Heritage Conservation in Macao”.
- T. Cunha Ferreira, P. M. Freitas, C. Frigolett, H. Mendonça, and A. T. Silva. 2024. The contribution of stakeholder engagement to cultural significance assessment: The case of values-based conservation management planning for the ocean swimming pool, portugal. *Built Heritage*, 8(1):26.
- Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. 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)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Leonardo (Don) A. N. Dioko, Patrick Lo, Wendy Tang, Virginia Hong, and Janice Leong. 2021. Macao visitor profile survey (vps). International Centre for Tourism and Hospitality Research (ICTHR), Macao University of Tourism (UTM). Online report page; results published annually; copyright © 2021.

- M. Foroughi, B. Andrade, A. Pereira Roders, and T. Wang. 2023a. Public participation and consensus-building in urban planning from the lens of heritage planning: A systematic literature review. *Cities*, 135:104235.
- M. Foroughi, B. de Andrade, and A. Pereira Roders. 2023b. Capturing public voices: The role of social media in heritage management. *Habitat International*, 142:102934.
- William C Gartner. 1994. Image formation process. *Journal of travel & tourism marketing*, 2(2-3):191–216.
- M. Ginzarly, A. Pereira Roders, and J. Teller. 2019. Mapping historic urban landscape values through social media. *Journal of Cultural Heritage*, 36:1–11.
- Patrick Goethals. 2016. Multilingualism and international tourism: a content-and discourse-based approach to language-related judgments in web 2.0 hotel reviews. *language and Intercultural communication*, 16(2):235–253.
- Durk Gorter. 2006. Introduction: The study of the linguistic landscape as a new approach to multilingualism. *International journal of multilingualism*, 3(1):1–6.
- Durk Gorter. 2013. Linguistic landscapes in a multilingual world. *Annual review of applied linguistics*, 33:190–212.
- Government of the Macao Special Administrative Region. 2026. Visitor arrivals for december and the whole year of 2025. Press release / news item.
- Brian Graham, Greg Ashworth, and John Tunbridge. 2000. *A geography of heritage*. Routledge.
- Xinxin Guo, Juho Pesonen, and Raija Komppula. 2021. Comparing online travel review platforms as destination image information agents. *Information Technology & Tourism*, 23(2):159–187.
- Scott A Hale. 2016. User reviews and language: how language influences ratings. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1208–1214.
- Germaine R Halegoua. 2020. The digital city: Media and the social production of place. In *The Digital City*. New York University Press.
- Nur Aliah Khairina Mohd Haris, Sofianita Mutalib, Ariff Md Ab Malik, Shuzlina Abdul-Rahman, and Siti Nur Kamaliah Kamarudin. 2023. Sentiment classification from reviews for tourism analytics. *International Journal of Advances in Intelligent Informatics*, 9(1):108–120.

- Rodney Harrison, Caitlin DeSilvey, Cornelius Holtorf, Sharon Macdonald, Nadia Bartolini, Esther Breithoff, Harald Fredheim, Antony Lyons, Sarah May, Jennie Morgan, et al. 2020. *Heritage futures: comparative approaches to natural and cultural heritage practices*. UCL press.
- Agnes Weiyun He. 2010. The heart of heritage: Sociocultural dimensions of heritage language learning. *Annual Review of Applied Linguistics*, 30:66–82.
- Songyao Huai and Tim Van de Voorde. 2022. Which environmental features contribute to positive and negative perceptions of urban parks? a cross-cultural comparison using online reviews and natural language processing methods. *Landscape and Urban Planning*, 218:104307.
- Assumpcio Huertas and Estela Marine-Roig. 2016. User reactions to destination brand contents in social media. *Information Technology & Tourism*, 15(4):291–315.
- ICOMOS. 2005. Advisory body evaluation: Historic centre of macao (china). Advisory Body Evaluation submitted to the World Heritage Committee.
- Bayan Issa, Muhammed Basheer Jasser, Hui Na Chua, and Muzaffar Hamzah. 2023. A comparative study on embedding models for keyword extraction using keybert method. In *2023 IEEE 13th International Conference on System Engineering and Technology (ICSET)*, pages 40–45. IEEE.
- Adam Jaworski, Nikolas Coupland, and Dariusz Galasinski. 2012. *Metalanguage: Social and ideological perspectives*, volume 11. Walter de Gruyter.
- Ingrid Jeacle and Chris Carter. 2011. In tripadvisor we trust: Rankings, calculative regimes and abstract systems. *Accounting, Organizations and Society*, 36(4-5):293–309.
- C. Johnston. 2017. Recognising connection: Social significance and heritage practice. *Córima, Revista de Investigación en Gestión Cultural*, 2(2).
- Jukka Jokilehto. 2008. The world heritage list: What is ouv? defining the outstanding universal value of cultural world heritage properties. Technical report, International Council on Monuments and Sites (ICOMOS). Monuments and Sites; accessed 2026-01-24.
- T. Katrakazis, A. Heritage, C. Dillon, P. Juvan, and S. Golfomitsou. 2018. Enhancing research impact in heritage conservation. *Studies in Conservation*. Advance online publication.

- Muhammad Qasim Khan, Abdul Shahid, M Irfan Uddin, Muhammad Roman, Abdullah Alharbi, Wael Alosaimi, Jameel Almalki, and Saeed M Alshahrani. 2022. Impact analysis of keyword extraction using contextual word embedding. *PeerJ Computer Science*, 8:e967.
- Andrei P Kirilenko, Svetlana O Stepchenkova, Hany Kim, and Xiang Li. 2018. Automated sentiment analysis in tourism: Comparison of approaches. *Journal of Travel Research*, 57(8):1012–1025.
- Olga Koblet and Ross S. Purves. 2020. From online texts to landscape character assessment: Collecting and analysing first-person landscape perception computationally. *Landscape and Urban Planning*.
- Philip Kotler. 2002. *Marketing places*. Simon and Schuster.
- N. I. Koufodontis and E. Gaki. 2022. UNESCO urban world heritage sites: Tourists’ awareness in the era of social media. *Cities*, 127:103744.
- S. Labadi. 2013. *UNESCO, Cultural Heritage, and Outstanding Universal Value: Value-Based Analyses of the World Heritage and Intangible Cultural Heritage Conventions*. AltaMira Press.
- Sofia Labadi and William Logan. 2015. Urban heritage, development and sustainability. *International frameworks, national and local governance*. London and New York: Routledge.
- Ivan Ka Wai Lai and Michael Hitchcock. 2017. Local reactions to mass tourism and community tourism development in macau. *Journal of Sustainable Tourism*, 25(4):451–470.
- Daniel Leung, Rob Law, and Hee Andy Lee. 2011. The perceived destination image of hong kong on ctrip. com. *International journal of tourism research*, 13(2):124–140.
- J. Li, S. Krishnamurthy, A. Pereira Roders, and P. van Wesemael. 2020. Community participation in cultural heritage management: A systematic literature review comparing chinese and international practices. *Cities*, 96:102476.
- Stephen W Litvin, Ronald E Goldsmith, and Bing Pan. 2008. Electronic word-of-mouth in hospitality and tourism management. *Tourism management*, 29(3):458–468.

- Yong Liu, Thorsten Teichert, Matti Rossi, Hongxiu Li, and Feng Hu. 2017. Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism management*, 59:554–563.
- K. Lynch. 1960. *The Image of the City*. MIT Press.
- Macao Government Tourism Office. n.d. Macao world heritage. Tourist brochure with map and site descriptions.
- Macao Special Administrative Region. 2013. Law no. 11/2013: Cultural heritage protection law. Boletim Oficial da Região Administrativa Especial de Macau.
- Estela Marine-Roig and Salvador Anton Clavé. 2016. A detailed method for destination image analysis using user-generated content. *Information Technology & Tourism*, 15(4):341–364.
- Estela Marine-Roig and Berta Ferrer-Rosell. 2018. Measuring the gap between projected and perceived destination images of catalonia using compositional analysis. *Tourism management*, 68:236–249.
- Juan Pedro Mellinas and Maria Sicilia. 2024. Comparing google reviews and tripadvisor to help researchers select the more appropriate information source. *Consumer Behavior in Tourism and Hospitality*, 19(4):646–655.
- Lynn Meskell. 2018. *A future in ruins: UNESCO, world heritage, and the dream of peace*. Oxford University Press.
- Elisa Oteros-Rozas, Berta Martín-López, Nora Fagerholm, Claudia Bieling, and Tobias Plieninger. 2018. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five european sites. *Ecological Indicators*, 94:74–86.
- Y. Pan, N. H. Nik Hashim, and H. C. Goh. 2024. Public perception of cultural ecosystem services in historic districts based on biterm topic model. *Scientific Reports*, 14:11717.
- N. D. Payntar. 2021. Archaeological heritage ‘on-the-rise’: Detecting emerging tourist attractions in peru using machine learning and geotagged photographs. *International Journal of Heritage Studies*, 28(3):222–244.
- John Pendlebury. 2013. Conservation values, the authorised heritage discourse and the conservation-planning assemblage. *International Journal of Heritage Studies*, 19(7):709–727.

- Francisco Vizeu Pinheiro, Koji Yagi, and Miki Korenaga. 2005. St. paul college: Historical role and influence in the development of macao. *Journal of Asian Architecture and Building Engineering*, 4(1):43–50.
- Telmo Pires, Eva Schlinger, and Dan Garrette. 2019. How multilingual is multilingual bert? *Preprint*, arXiv:1906.01502.
- Douglas Pocock and Ray Hudson. 1978. *Images of the urban environment*. Columbia University Press.
- I. Poullos. 2014. Discussing strategy in heritage conservation: Living heritage approach as an example of strategic innovation. *Journal of Cultural Heritage Management and Sustainable Development*, 4(1):16–34.
- Ross S Purves, Philipp Striedl, Inhye Kong, and Asifa Majid. 2023a. Conceptualizing landscapes through language: The role of native language and expertise in the representation of waterbody related terms. *Topics in Cognitive Science*, 15(3):560–583.
- Ross S. Purves et al. 2023b. Conceptualizing landscapes through language: The conceptualisation of waterbodies in german and english. *Topics in Cognitive Science*. Wiley landing page.
- Célia S. Rafael and Aurélia R. Almeida. 2017. Socio-demographic tourist profile and destination image in online environment. *Journal of Advanced Management Science*, 5(5):373–379.
- S Mostafa Rasoolimanesh, Mastura Jaafar, Ned Kock, and A Ghafar Ahmad. 2017. The effects of community factors on residents’ perceptions toward world heritage site inscription and sustainable tourism development. *Journal of Sustainable Tourism*, 25(2):198–216.
- Nils Reimers and Iryna Gurevych. 2019a. 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*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics.
- Nils Reimers and Iryna Gurevych. 2019b. Sentence-bert: Sentence embeddings using siamese bert-networks. *Preprint*, arXiv:1908.10084.
- Stuart Rose, Dave Engel, Nick Cramer, and Wendy Cowley. 2010. Automatic Keyword Extraction from Individual Documents. In *Text Mining: Applications and Theory*, pages 1–20, Chichester, UK. John Wiley & Sons.

- Francisco González Santa-Cruz and Tomás López-Guzmán. 2017. Culture, tourism and world heritage sites. *Tourism Management Perspectives*, 24:111–116.
- Karen Simonyan and Andrew Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. *Preprint*, arXiv:1409.1556.
- Laurajane Smith. 2006. *Uses of heritage*. Routledge.
- Laurajane Smith. 2020. *Emotional heritage: Visitor engagement at museums and heritage sites*. Routledge.
- State Party of the People’s Republic of China. 2005. Nomination of the historic centre of macao (china) for inscription on the world heritage list. Nomination file No. 1110 submitted to the World Heritage Centre.
- Svetlana Stepchenkova and Fangzi Zhan. 2013. Visual destination images of peru: Comparative content analysis of dmo and user-generated photography. *Tourism management*, 36:590–601.
- Bogusia Temple and Alys Young. 2004. Qualitative research and translation dilemmas. *Qualitative research*, 4(2):161–178.
- M. Tenzer and J. Schofield. 2024. Using topic modelling to reassess heritage values from a people-centred perspective: Applications from the north of england. *Cambridge Archaeological Journal*, 34(1):147–168.
- Yi-Fu Tuan. 1977. *Space and Place: The Perspective of Experience*. University of Minnesota Press, Minneapolis.
- UNESCO. 1972. Convention concerning the protection of the world cultural and natural heritage. Adopted by the General Conference at its seventeenth session, Paris, 16 November 1972.
- UNESCO. 2011. Recommendation on the historic urban landscape.
- UNESCO, ICCROM, ICOMOS, and IUCN. 2013. *Managing Cultural World Heritage*. UNESCO.
- UNESCO World Heritage Centre. 2005. Historic centre of macao. World Heritage List, Property No. 1110.
- UNESCO World Heritage Centre. 2014. Engaging local communities in stewardship of world heritage. World Heritage Papers Series No. 40.
- UNESCO World Heritage Centre. 2024. Operational guidelines for the implementation of the world heritage convention. WHC.23/01.

- John Urry. 1992. The tourist gaze “revisited”. *American behavioral scientist*, 36(2):172–186.
- Laurens van der Maaten, Geoffrey Hinton, and Yoesoep Rachmad. 2008. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9:2579–2605.
- Egbert Van der Zee, Nicola Camatti, Dario Bertocchi, and Khalid WA Shomali. 2024. Unesco world heritage site label and sustainable tourism in europe: a user-generated content analysis. *Regional Studies*, 58(10):1858–1873.
- José Van Dijck, Thomas Poell, and Martijn De Waal. 2018. *The platform society: Public values in a connective world*. Oxford university press.
- Eva Vanmassenhove, Dimitar Shterionov, and Andy Way. 2019. Lost in translation: Loss and decay of linguistic richness in machine translation. *arXiv preprint arXiv:1906.12068*.
- Loes Veldpaus, Ana R Pereira Roders, and Bernard JF Colenbrander. 2013. Urban heritage: putting the past into the future. *The Historic Environment: Policy & Practice*, 4(1):3–18.
- Ramya Venkateswaran, Robert Weibel, and Ross S Purves. 2014. Exploring and visualizing differences in geographic and linguistic web coverage. *Transactions in GIS*, 18(6):852–876.
- Flurina M. Wartmann et al. 2018. Investigating sense of place as a cultural ecosystem service in different landscapes through the lens of language. *Landscape and Urban Planning*. ScienceDirect landing page.
- Emma Waterton and Laura Jane Smith. 2010. The recognition and misrecognition of community heritage. *International journal of heritage studies*, 16(1-2):4–15.
- Zheng Xiang, Qianzhou Du, Yufeng Ma, and Weiguo Fan. 2017. A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism management*, 58:51–65.
- Zheng Xiang and Ulrike Gretzel. 2010. Role of social media in online travel information search. *Tourism management*, 31(2):179–188.
- Na Yang. 2025. Opinion sharing in online travel communities: A corpus-based comparison of members’ stance expression in ctrip and tripadvisor. *Discourse, Context & Media*, 65:100885.

- Benxiang Zeng and Rolf Gerritsen. 2014. What do we know about social media in tourism? a review. *Tourism management perspectives*, 10:27–36.
- X. Zhao, K. Xiong, and M. Zhang. 2025. Quantitative research on aesthetic value of the world heritage karst based on UGC data: A case study of huangguoshu scenic area. *PLOS ONE*, 20(2):e0317304.

A. Additional Tables

Name (TW)	Name (CN)	Name (EN)	POI ID	Lat	Long
媽祖閣 (媽閣廟)	妈祖庙 (妈阁庙)	A-Ma Temple	a_ma_temple	22.1864	113.5313
海事及水務局大樓	海事及水务局大楼	Moorish Barracks	moorish_barracks	22.1878	113.5327
亞婆井前地	亚婆井前地	Lilau Square	lilau_square	22.1884	113.5350
鄭家大屋	郑家大屋	Mandarin's House	mandarin's house	22.1888	113.5349
聖老楞佐堂及前地	圣老楞佐堂及前地	St. Lawrence's Church	st_lawrence	22.1907	113.5367
聖若瑟修院大樓	圣若瑟修院大楼	St. Joseph's Seminary	st_joseph	22.1918	113.5373
崗頂前地	岗顶前地	St. Augustine's Square	st_augustine	22.1923	113.5378
崗頂劇院	岗顶剧院	Dom Pedro V Theatre	dom_pedro	22.1920	113.5384
何東圖書館大樓	何东图书馆大楼	Sir Robert Ho Tung	ho_tung_lib	22.1926	113.5373
聖奧斯定教堂	圣奥斯特定教堂	St. Augustine's Church	st_augustine	22.1922	113.5384
市政署大樓	市政署大楼	Municipal Affairs Bureau	municipal_aff	22.1933	113.5395
議事亭前地	议事亭前地	Senado Square	senado_sq	22.1936	113.5398
三街會館 (關帝廟)	三街会馆	Sam Kai Vui Kun	sam_kai_vui	22.1941	113.5394
仁慈堂大樓	仁慈堂大楼	Holy House of Mercy	holy_house	22.1938	113.5402
主教座堂	主教座堂	Cathedral	cathedral	22.1936	113.5414
盧家大屋	卢家大屋	Lou Kau Mansion	lou_kau	22.1943	113.5413
玫瑰聖母堂 (板樟堂)	玫瑰圣母堂	St. Dominic's Church	st_dominic	22.1947	113.5404
聖保祿學院址	圣保禄学院遗址	Ruins of St. Paul's	ruins_st_paul	22.1976	113.5409
哪吒廟	哪吒庙	Na Tcha Temple	na_tcha_temp	22.1978	113.5406
城牆遺跡	城墙遗迹	Old City Walls	old_city_wall	22.1978	113.5406
大炮臺	大炮台	Mount Fortress	mount_fort	22.1969	113.5423
聖安多尼堂及前地	圣安多尼堂及前地	St. Anthony's Church	st_anthony	22.1991	113.5394
東方基金會會址	东方基金会会址	Orient Foundation	orient_found	22.2004	113.5399
基督教墳場	基督教坟场	Protestant Cemetery	protestant_cem	22.1998	113.5398
東望洋炮台、燈塔	东望洋炮台	Guia Fortress	guia_fortress	22.1967	113.5497
媽閣廟前地	妈阁庙前地	Barra Square	barra_square	22.1867	113.5307
大堂前地	大堂前地	Cathedral Square	cathedral_sq	22.1936	113.5414
板樟堂前地	板樟堂前地	St. Dominic's Square	st_dominic_sq	22.1946	113.5405
白鴿巢前地	白鸽巢前地	Camoës Square	camoes_square	22.1994	113.5396
耶穌會紀念廣場	耶稣会纪念广场	Company of Jesus Square	jesus_sq	22.1968	113.5408

Table A.1.: Heritage landmarks with multilingual names and coordinates.

Heritage Landmark	Ctrip	GM EN	GM SC	GM TC	TA	Img EN	Img SC	Img TC
A-Ma Temple	100	100	100	100	100	257	216	318
Cathedral	30	100	100	100	82	207	267	180
Cathedral Square	0	35	35	29	27	77	77	77
Dom Pedro V Theatre	13	100	100	99	22	251	258	227
Guia Fortress	67	100	100	100	100	233	292	250
Sir Robert Ho Tung Library	21	78	80	78	25	175	179	179
Holy House of Mercy	38	70	70	67	21	120	120	120
Company of Jesus Square	0	9	9	6	7	25	25	23
Lilau Square	35	86	84	84	29	174	164	170
Lou Kau Mansion	39	100	100	100	49	237	292	254
Mandarin's House	50	100	100	100	100	171	272	315
Moorish Barracks	10	53	55	53	16	85	90	90
Mount Fortress	100	100	100	100	100	308	215	276
Municipal Affairs Bureau	73	100	100	99	76	231	251	194
Na Tcha Temple	54	99	97	96	42	174	184	176
Orient Foundation	0	41	41	37	6	61	61	61
Protestant Cemetery	0	24	24	24	15	45	45	45
Ruins of Saint Paul's	101	100	100	100	100	338	257	316
Sam Kai Vui Kun	11	50	50	47	10	82	82	82
Senado Square	100	100	100	100	100	347	174	309
St. Anthony's Church	19	88	99	100	23	154	177	169
St. Augustine's Church	14	80	79	73	18	174	170	174
St. Augustine's Square	14	66	65	59	49	112	110	116
St. Dominic's Church	100	100	100	100	100	237	213	266
St. Dominic's Square	0	9	9	9	43	27	27	27
St. Joseph's Seminary	15	99	91	89	15	155	173	170
St. Lawrence's Church	22	100	100	100	52	215	229	203

Table A.2.: Review and image counts per heritage landmark. Ctrip = Ctrip reviews; GM = Google Maps reviews; TA = TripAdvisor reviews; Img = image-based analysis. Language codes denote English (EN), Simplified Chinese (SC), and Traditional Chinese (TC).

ID	Crit.	Claim (English)
2.1	ii	Macao's strategic coastal position and the long-term interchange of cultural, scientific, technological, artistic and architectural values between Chinese and Western actors.
2.2	ii	Macao's strategic position on the Chinese coast made it a key node in global maritime trade routes between China and the wider world.
2.3	ii	The special relationship negotiated between Chinese authorities and the Portuguese allowed a long-term, regulated foreign settlement in Macao.
2.4	ii	Macao served as an important hub where people, ideas and technologies circulated between China, Europe and other parts of Asia.
2.5	ii	Over several centuries, Macao became a place where cultural, scientific, technological, artistic and architectural ideas were exchanged between Chinese and Western actors.
2.6	ii	Institutions such as theatres, schools, churches, hospitals and fortifications in Macao embody this continuous transfer of knowledge and techniques.
3.1	iii	Macao as the first and longest-lasting European settlement in China, bearing unique testimony to a sustained encounter between the West and China visible in its mixed culture and historic core.
3.2	iii	Macao is the first and longest-lasting European settlement in China, representing an enduring encounter between the West and China from the 16th to the 20th century.
3.3	iii	For centuries Macao was a focal point for traders, missionaries and scholars who used the city as a base to work in China and wider Asia.
3.4	iii	The historic core shows a visible fusion of Western and Chinese architectural styles, building techniques and decorative motifs.
3.5	iii	Macao's urban landscape, with churches, temples and civic buildings standing side by side, testifies to a mixed and layered cultural identity.
3.6	iii	Local language, customs and everyday life in Macao developed under both Chinese and Western influences, creating a distinctive multicultural community.
4.1	iv	Macao as an outstanding architectural ensemble and urban fabric that illustrates four and a half centuries of interaction between Western and Chinese civilisations along a historical route from harbour to city.

Continued on next page

ID	Crit.	Claim (English)
4.2	iv	Macao's historic centre forms an architectural ensemble that documents four and a half centuries of interaction between Western and Chinese building traditions.
4.3	iv	A historical route runs from the ancient Chinese harbour and A-Ma Temple to the former Christian city, lined with squares, churches, temples and civic buildings that illustrate this encounter.
4.4	iv	The dense street network, with narrow, winding lanes opening onto public squares, reflects typical European port-city patterns adapted to a Chinese setting.
4.5	iv	Portuguese and Chinese architectural elements are often blended within single buildings, such as Western churches with Chinese decorative details or layouts.
4.6	iv	Visual connections between the city and the surrounding seascape, including the Inner Harbour and hilltop fortresses, recall Macao's origin and function as a coastal trading port.
6.1	vi	Macao's role as a channel for cultural, spiritual, scientific and technical influences whose ideas contributed to major changes in China, including the end of the imperial system and the rise of the republic.
6.2	vi	Macao was a key channel for transmitting cultural, religious and intellectual ideas between Western societies and China.
6.3	vi	Missionaries and reformers used Macao as a base to introduce new educational, medical and charitable practices into Chinese society.
6.4	vi	Several first Western-style institutions in China, such as an early theatre, college and modern hospital, were established in Macao.
6.5	vi	The city provided an environment where Chinese intellectuals, including Sun Yat-sen, encountered new political and social ideas that influenced the end of the imperial system and the creation of the Republic.
6.6	vi	Religious practices, festivals and charitable organisations in Macao reflect a long-term blending of Christian and Chinese spiritual traditions.

Table A.3.: Claim-level statements extracted from UNESCO and ICOMOS documentation for the Historic Centre of Macao, organised by cultural criterion.

Component	Sentences	Sent. mean	Top keyterms	CII	CIII	CIV	CVI	Dominant theme
Ruins of Saint Paul's	1166	0.45	temple; antique style; known as	7.38%	26.42%	64.75%	1.46%	Buildings & Streetscapes
A-Ma Temple	1081	0.47	church; nativity; ruins	4.72%	14.52%	65.77%	14.99%	Green Spaces & Parks
St. Dominic's Square	994	0.47	square; landmarks; cross	3.02%	22.13%	66.10%	8.75%	Catholic Church Interiors
Headquarters Building of the Municipal Affairs Bureau (Former Leal Senado Building)	883	0.45	square; ruins; festival	10.76%	30.46%	49.60%	9.17%	Interiors & Decorative Details
Senado Square	874	0.47	theater; cultural; balcony	14.07%	23.80%	56.06%	6.06%	Buildings & Streetscapes
Mount Fortress	858	0.46	lighthouse; church; fortress	9.09%	18.18%	70.75%	1.98%	Green Spaces & Parks
Mandarin's House	837	0.50	catholic; architectural; elderly home	4.90%	36.44%	51.02%	7.65%	Interiors & Decorative Details
Lou Kau Mansion	785	0.45	municipal; Portugal; garden	6.50%	44.33%	44.71%	4.46%	Interiors & Decorative Details
Na Tcha Temple	772	0.42	garden; gallery; mansion	2.59%	17.62%	72.41%	7.38%	Temples & Folk Shrines
St. Lawrence's Church and forecourt	738	0.47	mountain spring; fountain; Portugal	5.28%	18.43%	67.75%	8.54%	Catholic Church Interiors
Guia Fortress	697	0.48	residence; courtyard; stucco relief	8.75%	23.10%	65.28%	2.87%	Green Spaces & Parks
Cathedral	682	0.48	church; nativity; ruins	3.08%	17.60%	69.65%	9.68%	Facades
St. Anthony's Church and forecourt	651	0.43	historic; antique style; old house	3.07%	12.29%	75.12%	9.52%	Catholic Church Interiors
Dom Pedro V Theatre	537	0.45	fortress; climb; St.Paul	10.80%	28.31%	43.76%	17.13%	Interiors & Decorative Details
St. Joseph's Seminary Church, forecourt and staircase	499	0.48	incense; St.Paul; temple	4.61%	14.83%	65.93%	14.63%	Catholic Church Interiors
St. Augustine's Church	494	0.43	cemetery; tombstone; bible	4.05%	14.37%	73.48%	8.10%	Catholic Church Interiors
Headquarters Building of the Holy House of Mercy of Macao	477	0.43	church; ruins; crowds	5.03%	25.79%	53.25%	15.93%	Buildings & Streetscapes
Lilau Square	444	0.43	temple; guildhall; China	14.19%	19.37%	63.74%	2.70%	Green Spaces & Parks
Sir Robert Ho Tung Library	418	0.43	fountain; plaza; shopping	16.03%	27.51%	40.19%	16.27%	Green Spaces & Parks
St. Augustine's Square	345	0.47	garden; library; historic	10.72%	17.68%	67.83%	3.77%	Facades
Sam Kai Vui Kun	342	0.41	church; heritage; building	4.39%	19.01%	61.11%	15.50%	Temples & Folk Shrines
Moorish Barracks	253	0.41	church; statue; plaza	5.53%	26.48%	62.06%	5.93%	Lanes, Alleys & Walls
Protestant Cemetery	223	0.44	theater; library; plaza	4.04%	26.01%	61.88%	8.07%	Green Spaces & Parks
Cathedral Square	215	0.48	church; baroque; ruins	4.65%	13.02%	76.28%	6.05%	Buildings & Streetscapes
Headquarters Building of the Orient Foundation	205	0.42	church; square; building	16.59%	35.12%	43.41%	4.88%	Green Spaces & Parks
St. Dominic's Church	138	0.51	church; antique style; building	5.80%	16.67%	74.64%	2.90%	Buildings & Streetscapes
Company of Jesus Square	38	0.44	church; baroque; Portugal	13.16%	15.79%	65.79%	5.26%	Buildings & Streetscapes

Table A.4.: Component-level summary of textual volume (sentences), mean sentiment, top keyterms, Stage-2 Macao-specific OUV criterion composition (II/III/IV/VI), and dominant visual theme. Percentages may not sum to 100 due to rounding.

Personal Declaration

I declare that the submitted thesis is the result of my independent work. All external sources are explicitly acknowledged in the thesis.

I further declare that the following tools were used in the development of this work: OpenAI's ChatGPT was used as a support tool for brainstorming, outlining, and troubleshooting code and document preparation (e.g., Python, LaTeX, and data-processing scripts). Google Translate was used to assist with the translation and standardisation of multilingual key terms during exploratory stages.

I am aware that I bear full responsibility for the selection, verification, and integration of any outputs produced with these tools, and for all results and interpretations presented in the thesis. I declare that the use of generative AI has been fully disclosed in accordance with the applicable guidelines, and that any external sources and tools are explicitly acknowledged in the thesis.

Yilin Wei
Signature

Zürich, 26.01.2026
Place, date

Full name: Yilin Wei

Matriculation number: 23-742-570