Masters
Program
in **Geospatial Technologies**

Vernacular boundaries of historic neighborhoods in the city of Lisbon

Mónica Sofía Roncancio Blanco

Dissertation submitted in partial fulfilment of the requirements for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS Information Management School
UNIVERSITAT JAUME·I
Westfälische Wilhelms-Universität Münster

Vernacular Boundaries of Historic Neighborhoods in the City of Lisbon

Dissertation supervised by

Prof. Dr. Marco Octávio Trindade Painho, PhD
NOVA Information Management School
Lisbon, Portugal

Co-supervised by:

Vicente De Azevedo Tang
NOVA Information Management School
Lisbon, Portugal

Prof. Dr. Sven Casteleyn, PhD
Universitat Jaume I
Castelló, Spain

February, 2024

Declaration of Originality

I declare that the work described in this document is my own and not from someone else. All the assistance I have received from other people is duly acknowledged and all the sources (published or not published) are referenced.

This work has not been previously evaluated or submitted to NOVA Information Management School or elsewhere.

Lisbon, February 26, 2024

Mónica Sofía Roncancio Blanco

[the signed original has been archived by the NOVA IMS services]

Acknowledgments

First, I extend my deepest gratitude to Prof. Dr. Marco Painho for his support and guidance throughout this academic journey. He has been an incredible professor and human being.

I am also thankful to Dr. Sven Casteleyn and Vicente De Azevedo Tang for their insights and suggestions as my thesis supervisors. Their patience and guidance have been invaluable. Additionally, My beloved friend Nyi Nyi Nyan Lin deserves special mention for his invaluable support over these 6 months, offering always accurate recommendations to improve my project.

This program has afforded me the opportunity to fulfill my dream of studying abroad, allowing me to be part of three European Universities and to establish academic, professional, and personal connections.

Finally, my heartfelt thanks go to my family and friends for their daily support from afar. This new academic and personal achievement is undoubtedly for them.

This thesis is one of the outputs of a project funded by the Fundação para a Ciência e a Tecnologia (FCT), named "CityMe," with the Project Reference (EXPL/GESURB/1429/2021).

Vernacular boundaries of historic neighborhoods in the city of Lisbon

Abstract

This research investigates the delineation of vernacular boundaries in Lisbon's historic neighborhoods: Alfama, Mouraria, and Bairro Alto, the study explores the integration of perceived boundaries derived from residents' sketches and geo-tagged data from online activities. Utilizing a two-phase methodology, the research first extracts perceived boundaries through web-based surveys categorized by residents' length of stay, distinguishing between short-term (less than 10 years) and long-term residents (more than 10 years). This approach allows for an overlay analysis, identifying Core and Domain regions based on consensus thresholds. Secondly, the study retrieves geo-tagged boundaries using A-DBSCAN and alpha-shape algorithms to analyze online activity, offering a comparative analysis with the perceived boundaries.

The findings reveal a nuanced understanding of how residents and online users conceptualize neighborhood spaces, highlighting discrepancies and convergences between perceived and digital mappings. By calculating the Intersection Over Union (IOU) and F-scores, the research quantitatively assesses the overlap between different data sources, identifying the most accurate delineations that reflect the historic neighborhoods' spatial reality. This study contributes to urban planning and policymaking by providing insights into residents' spatial perceptions, emphasizing the importance of considering both lived experiences and digital footprints in the mapping of urban areas. The research underscores the potential of combining traditional survey methods with innovative geo-spatial technologies to enhance the precision and relevance of urban geographic studies.

Sustainable Development Goals (SGD):

Keywords

Vernacular Boundaries

Perceived Boundaries

User Generated Content

Clustering Analysis

Geographic Information Science

Spatial analysis

Historic neighborhoods

Acronyms

A-DBSCAN – Approximate Density-Based Spatial Clustering of Applications with Noise

GIS – Geographic Information Systems

IOU – the Intersection Over Union

NBHD –Neighborhood

UGC – Use Generated Content

Index of the Text

Index of Tables

Index of Figures

Appendix

Chapter 1

1. Introduction

Lisbon experienced an astounding 132.6% increase in overnight stays in 2022 (Vernon, 2023). This growth continued into 2023, as confirmed by Eurostat ('Tourism Industry Eclipses Pre-Pandemic Levels in 2023,' Eurostat, 2023), signifying the city's rising popularity among global urban holiday enthusiasts. This trend underscores Lisbon's dynamic response to intense global competition, as cities vie for a stronger position in the international tourist market (Baptista et al., 2018). Over the years, Lisbon has undergone significant economic, social, and cultural transformations, driven by policies since the 1990s aimed at attracting tourists and foreign investment, particularly in its flexible real estate sector (Cocola-Gant & Gago, 2021); However, this success has not come without challenges. The city is facing a severe housing crisis, exacerbated by tourism and foreign investment strategies, leading to gentrification and over-tourism, especially in historic neighborhoods (Da Costa, 2008). These areas struggle with defining their vernacular boundaries, highlighting the intricate nature of urban transformations.

The investigation of vernacular boundaries, integrating insights from sociology, psychology, geography, and community studies (Chugunov et al., 2019), tackles the issue of data subjectivity, which is deeply influenced by personal perceptions and diverse definitions. The study of vernacular geography within historical neighborhoods grapples with the task of accurately mapping areas imbued with strong cultural connections, which often result in spatially diffuse and complex expansions. This complexity is compounded by the fact that official administrative divisions frequently fail to mirror the actual spatial and social dimensions of neighborhoods, a discrepancy especially pronounced within Lisbon's “Freguesias” (Hollenstein & Purves, 2010).

This research is motivated by the need to develop a nuanced understanding of Lisbon's historic neighborhoods beyond the limitations of conventional administrative delineations. It seeks to refine the approximation of vernacular boundaries that more accurately reflect the neighborhoods' spatial extents. Through the comparative analysis of two distinct data sources online web-based surveys and user-generated content (UGC) this study aims to make a significant contribution to GIScience. It addresses the challenge of the inherent imprecision in mapping historical neighborhoods and seeks to overcome the obstacles in capturing their true essence (Deng, 2016). In doing so, it aspires to offer insights that are critically valuable for the realms of urban planning and policymaking, ensuring that interventions and policies are deeply informed by the lived realities and cultural fabric of the communities they affect.

1.1 Aim and Research Questions

This project aims to extract, assess, and compare the spatial definition of three historic neighborhoods in Lisbon, using data from cognitive mapping and geo-tagged user-generated content. This leads us to address the following research questions:

- How do residents spatially define different historic neighborhoods in the city of Lisbon?
- What are the spatial footprints of historic neighborhoods according to different sources of geo-tagged user-generated content?
- What are the similarities and differences regarding neighborhood boundaries between primary and secondary data?
- How reliable is the use of geo-tagged online activity in inferring residents' opinions on neighborhood boundaries?

1.2 Project Objectives

This study aims to achieve the following objectives to address the posed research questions, applying tools and methods associated with the geographic information field:

- Extract representative boundaries of the historic neighborhoods from participant sketches collected using a web-based map survey.
- Retrieve representative boundaries from different sources of geo-tagged online activity.
- Quantitively compare the spatial definitions extracted from the cognitive mapping with boundaries from different sources of geo-tagged online activity for the historic neighborhoods.

1.3 Approach to Achieving the Project Objectives

To achieve these objectives, the research employs a comprehensive methodology in three phases as follows.

First Phase: Perceived boundaries are extracted through participant sketches collected via a web-based map survey. This stage aims to capture residents' perceptions of their neighborhood boundaries. Participants are categorized based on their length of residency in the city, distinguishing between short-term (less than 10 years) and long-term residents (more than 10 years). This classification facilitates an overlay analysis to perform Perceived Boundaries Extraction, yielding two consensus threshold-based regions: the Core and Domain regions.

Second Phase: Representative boundaries are retrieved from various sources of geo-tagged online activity, using algorithms such as A-DBSCAN and alpha-shape to calculate and define representative clusters and shapes.

Third Phase: This phase involves a comparative analysis between the perceived boundaries identified by the resident groups and the geo-tagged boundaries. The goal is to determine the congruence between residents' perceptions and digital representations of the historic neighborhoods. This comparison utilizes the Intersection Over Union (IOU) metric to quantify the extent of overlap between shapes, including the calculation of overlapping areas and F-scores for each dataset. The aim is to identify the most approximate delineations that accurately reflect the historic neighborhoods, providing valuable insights for urban planning and policymaking.

1.4 Thesis Organization

In the following sections: **Chapter 2** reviews the existing literature on the subject, setting the stage for the study by identifying gaps and framing the research questions within the context of current knowledge. **Chapter 3** describes the two-phase methodology employed in this research, which includes extracting perceived boundaries through web-based surveys and analyzing geo-tagged data from online activities. **Chapter 4** presents and discusses the comparative analysis between perceived and geo-tagged boundaries, highlighting the key findings and their implications for the study of urban spaces. **Chapter 5** addresses the research questions, providing detailed answers based on the analysis conducted in the previous chapters. **Chapter 6** summarizes the key findings and contributions of the study, offering a concise overview of the research outcomes and their relevance to the field.

Chapter 2

2.Literature Review

2.1 Studying Neighborhoods, Vernacular boundaries, and Place.

2.1.1 Neighborhoods

Neighborhoods, within a sociological context, are acknowledged as residential zones characterized by significant social and cultural compositions. They are places where residents share similar characteristics, driven by a collective struggle for urban space and resources. This conceptualization of neighborhoods represents one of the initial attempts to define these areas. It was proposed by urban sociologists Park and Burgess in 1925. Their seminal work, “The City,” laid the foundation for early 20th-century urban planning studies, marking a significant moment in this field, (Park et al.,1925).

In the following years, contributions like those from urban planner Kevin Lynch would highlight five key elements defining the city's image, drawn from research on mental mapping and collective perceptions. Among these elements, "districts" and "edges" are closely tied to the concept of neighborhood structure. Edges represent actual or perceived boundaries, delineated as linear features demarcating the peripheries of areas known as districts (Lynch, 1964). These districts vary in size from medium to large and showcase a diversity of characteristics, including types of buildings, resident demographics, topography, activities, historical significance, and levels of upkeep.

Currently, relevance in the depiction of these residential areas lies in the fact that neighborhoods boost social life, access to public services, and surveillance, connecting and exchanging resources with other neighborhoods (Bae & Montello, 2018). On the other hand, the neighborhood conformation is tied to the residents' interactions, which are essential to maintaining their vibration and diversity through work, housing, and recreational activities. According to Coulton study, neighborhoods are a collective and geographical construction, although these dimensions are rooted in social and psychological conditions, they are bound to geographic space and as a result, obtaining meaningful spatial representations of neighborhoods is necessary for the enactment of local policies (Coulton et al., 2013).

2.1.2 Vernacular Boundaries

Due to the existence of unclear borderlines, vernacular regions are distinguished by strong cultural linkages that produce spatial expansions that are not sharp, such as city center or historic neighborhoods. Some writers, such as W. Zelinsky (Zelinsky, 1980), investigated how people perceive space and how place awareness or regional consciousness creates an attachment sense based on popular culture. One of his most notable projects was the identification of cultural regions in North America by analyzing the frequency of local place names and businesses.

Scholars have found practical implications in the analysis of vernacular regions because official administrative units or census blocks do not necessarily represent neighborhood or region extension as spatial and social entities (Hollenstein & Purves, 2010). Consequently, administrative artefacts are overlain in a variety of factors associated to collective or individual people behaviors between social, economic, and environmental features such as housing systems, land use, and accessibility and whose resident's perception state remains in constant change thus vernacular studies going focus on the factors' extension, contributing to policymaking.

For Evans (Evans & Waters, 2007), we live daily in vernacular regions that we accentuate with geographical terms that are not formally represented. For example, "High crime areas". The terms act as a tool to understand the networks of sociolinguistic communities with shared understandings; while for Coulton, the vernacular regions scale is another feature of the study whose dwellers perception can be smaller or bigger according to their own experiences (Coulton et al., 2013).

Incorporating vernacular regions into research introduces significant challenges due to the subjective nature of data derived from individual and community perceptions, leading to varied definitions and interpretations. These challenges include difficulties in data collection, where the diversity in personal descriptions of surroundings can result in inconsistencies when mapped against official geographic boundaries, Additionally, the reliability of subjective perceptions raises questions, as personal biases and experiences may heavily influence the delineation of vernacular areas, (Deng, 2016). To address these issues, innovative methodologies are required for harmonizing vernacular boundaries with official geographic data, necessitating the development of information systems that prioritize lay user perspectives over traditional administrative geography. Such systems employ techniques like crowdsourced mapping and participatory

GIS to integrate vernacular areas into standard datasets, bridging the gap between subjective perceptions and objective data.

2.1.3 Places

The contribution in the definition of "Place" is wide due to fields such as geography, psychology, and anthropology among others being involved. The unique identity and significance seem to be remarkable features of place definition, according to the geographer Edward Relph in his book "Place and Placelessness" he claims his research method is "a phenomenology of place" that relies on the interpretative study of human experience. For him, a place is a geographical area with a distinct identity and significance to people. Human experiences, memories, and cultural importance are all present in places. They foster a feeling of identification and attachment to people and communities.

Belonging a place shapes people's personalities creating deep connections with the sense of self, places influence how communities form, how people relate to each other, and how social norms are established; on the other hand, cultural and historical meaning are content in these places displaying stories, cultural heritage, social memories and tiding with environmental relationships.

As an example of the aforementioned. Two authors Hernandez's research on place identity and Anacta et al.' study on spatial representation through route instructions contribute to our understanding of how individuals connect with and define their neighborhoods. Hernandez's work delves into the emotional bonds people form with their environment, measuring the intensity of place attachment and place identity through questionnaires,(Hernández et al., 2007). These questionnaires assess various environmental scales and neighborhood, city, and island to measure the 'type of bond' that individuals have with each. This approach underscores the significance of place in shaping personal identity and social dynamics. On the other hand, Anacta approach the concept of neighborhood from a spatial perspective, exploring how people cognitively map areas with indistinct boundaries when giving route directions, (Anacta et al., 2017). They focus on 'Neighborhood' as a category within sketch mapping, highlighting the social use of shared facilities and the homogeneity of residential or structural characteristics as defining elements. By combining the findings from both studies, we gain a comprehensive view of how place identity and spatial cognition interplay in neighborhoods' conceptualization and lived experience, enhancing our insights into urban social geography and the psychology of space.

2.2 Data Sources of Perceived Boundaries in Neighborhood Mapping

In the field of neighborhood mapping, it is essential to address the concept of cognitive regions informal areas characterized by unclear boundaries, membership status, extension, location, and shape, which are influenced by the distribution of communities or the perceptions of residents. These regions' spatial properties are also shaped by cultural factors, adding to their complexity. A comprehensive understanding of socio-demographics and environmental characteristics is vital for achieving accurate spatial delineation. This process involves precisely defining the boundaries and features of these areas, considering their informal nature and the impact of cultural and social elements.

Cognitive regions present significant challenges in geographic information science, mainly due to their vagueness, the difficulty in representing them accurately in GIS tools, their varying scales, and their lack of homogeneity or consistency. To capture the diverse perceptions and experiences of residents, several data sources are used, including surveys and interviews that verify the interviewees' defined characterization, participatory mapping exercises, analysis of social media, and examination of historical records. Additionally, preexisting administrative (notably census) data and other agency-defined zones serve as proxies for neighborhood delineations.

This project employs two of these data sources, which will be explained in detail below.

2.2.1 Online Web-Based Survey

Surveys have been a traditional method for comprehending cognitive boundaries. They involve directly querying local inhabitants about their perceptions of neighborhood limits. In advancing the mapping of diffuse areas, surveys have been utilized not only in delineating neighborhood areas but also in addressing issues affecting neighborhoods. For instance, the study by Evans & Waters employed a survey to assess the high levels of crime in LA city, identifying areas with a high risk of crime. This approach allowed researchers to gather nuanced insights into how residents mentally map their surroundings, (Evans & Waters, 2007). The study by Westerholt et al., used a web mapping-based survey that let respondents map and rate areas in the city related to points of interest extracted from Google Places, showing how individuals' sense of place influences their perception of neighborhood boundaries, (Westerholt et al., 2022).

2.2.2 User-Generated Content (UGC)

UGC is another rich data source for studying human behavior, sentiment, and perception of the urban environment. Social media analysis provides a modern, data-rich source for exploring perceived boundaries. The study by Hollenstein & Purves demonstrates how user-generated content on platforms like Flickr can offer insights into how people perceive and interact with urban spaces. This analysis allows for the extraction of cognitive boundaries based on the aggregation of individual experiences and expressions captured through social media, (Hollenstein & Purves, 2010).

Moreover, the work of Huang emphasizes the potential of combining "big data," represented by geotags and images from Instagram and Twitter, with "small data," such as questionnaires and sketch maps. Their study underlines the importance of analyzing social media content to map the cognitive landscapes of cities, revealing how residents and visitors perceive them. Findings suggest that social media analytics can provide a reliable measure of perceived city images, with platforms like Instagram explaining public perception associated with tourist attractions and landmarks, while Twitter is linked to the place's relevance to everyday life venues for communities. These results have practical implications, especially in urban planning, offering valuable information to policymakers and citizens alike, (Huang et al., 2021).

2. 3 Methods for Boundaries Extraction

2.3.1 Perception-based neighborhood delineation

Its method involves gathering unique insights through mental mapping, highlighting the cognitive mapping of neighborhoods as described by (Dalton & Hurrell, 2023). These mappings are defined as graphic and subjective representations that capture not just the physical layout but also the emotional and cultural associations individuals hold with their communities. This approach measures cognitive representations by sampling residents, collecting socio-demographic features, and retrieving mental representations through surveys or interviews. The collected data then facilitates spatial analysis about core areas. The level of agreement among the mental maps serves as a measure of membership, revealing collective perceptions of neighborhood boundaries.

2.3.2 Boundaries Based on Online Activity

(Deng, 2016) discussed the interactions among residents and how these social processes relate to geographic patterns and neighborhood formation, focusing on physical networks. The method gains relevance with the incorporation of User-Generated Content (UGC), as it characterizes a neighborhood and reveals details of its social composition and environment, transforming geographical data into neighborhood patterns. A case study in Spain utilized data from social media platforms like Google Places, Foursquare, Twitter, and Instagram. Previous research has shown the complementarity of these four sources for inferring valuable insights regarding the spatiotemporal use of city spaces and people's preferences, aiming to leverage the availability of user-generated data sources (Bernabeu-Bautista et al., 2023). Additionally, studies by (Gao et al., 2017a) and (Tang et al., 2022) use user-generated content to determine the delimitations of cognitive regions. They perform cluster analysis using algorithms such as A-DBSCAN to identify point clusters based on the collected data, aiming to find core areas that represent a consensus among a significant number of users who agree that these areas represent regional or neighborhood extensions.

2.4 The Study Site

2.4.1 Lisboa

Lisbon, the capital of Portugal, is a dynamic city with a rich history. It started as a Phoenician trade post, was taken over by the Romans and then the Muslims and became the departure port for the famous Portuguese explorers. It's one of the oldest cities in Europe. Nowadays, it's a popular destination for international students, tourists, and foreign residents (Tang et al., 2021).

Each phase of Lisbon's development resulted in various settlements, now recognized as the historic neighborhoods in the city center. For this study, three neighborhoods, Alfama, Mouraria, and Bairro Alto, have been selected. These areas, significant for their historic importance and the need for clear delineation, are situated within diverse administrative boundaries Figure 1. Some of the main characteristics of each historic neighborhood are depicted in the next section.

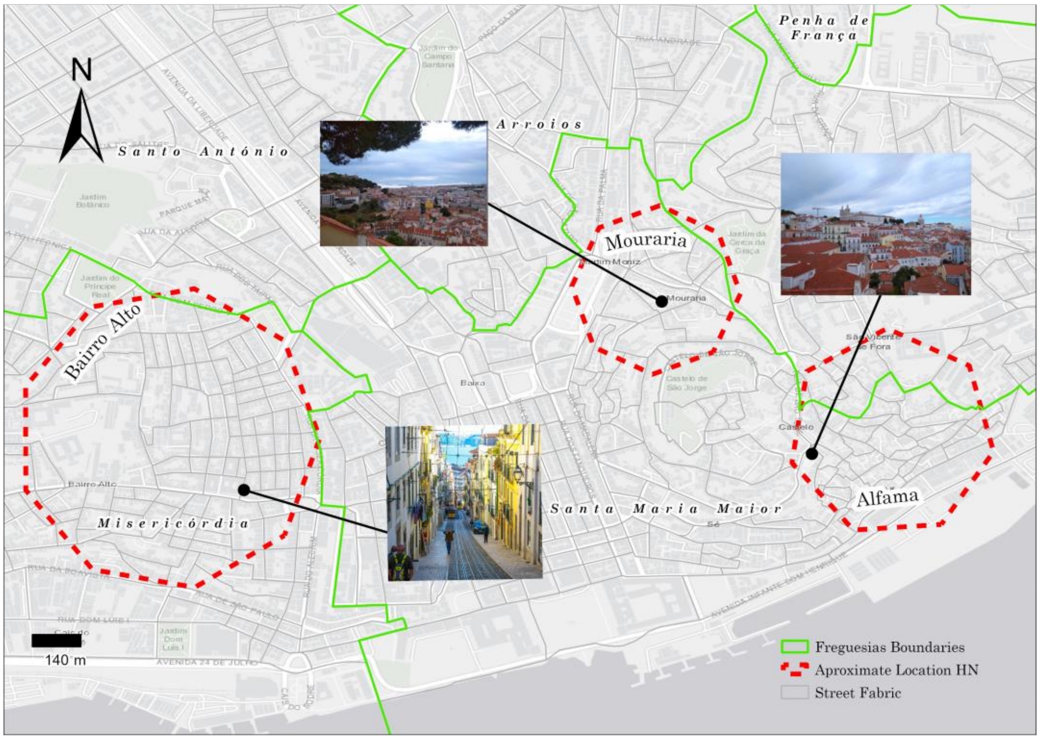

Figure 1. Freguesías Boundaries and Perceived Historic Neighbourhoods location map

2.4.2 Alfama

Named after the Arabic al-hamma. The hill where Alfama is located, along with the nearby castle hill, were among the first settled areas that contributed to Lisbon's origins. The part of Alfama stretches along the hill between the Church of São Miguel and Santo Estêvão, with Rua da Regueira at its core. The neighborhood extends up to where the Escolas Gerais used to be, a residence for university students until the 16th century.

At the hill's lower end lies the shore of the Tejo River, once a center for river commerce and a fishing neighborhood during the 13th and 14th centuries.

Alfama has always been a tourist attraction due to being the oldest part of the city, Moorish-rooted architecture that preserves the ancient medieval Arab road mall with alleyways, narrow streets and stairs that make it famous as well as its proximity to the São Jorge castle and its connection with Fado, the traditional Portuguese folk music. This neighborhood has an important social visibility that is widely recognized in aspects such as literature, painting, and music in the city of Lisbon, becoming a historical heritage spot (Da Costa, 2008).

In past centuries, Alfama was home to both aristocrats and working-class individuals. The diverse backgrounds of its residents fostered a strong sense of community, rooted in long-term tenancy and the social connections

formed among neighbors (Da Costa, 2008). However, recent challenges such as over-tourism, the rise in lodging phenomena, and damage to properties indicate that Alfama is undergoing significant transformations, gradually morphing into a predominantly tourist district (Cocola-Gant & Gago, 2021).
At present, Alfama is renowned as Lisbon's most picturesque neighborhood, often seen as a journey back in time. It's a popular spot-on Instagram, frequently showcased by tourists, ERASMUS students, and temporary residents, (Baptista et al., 2018).

2.4.3 Mouraria

Mouraria neighborhood is located between the valley of the Plaça Martin Moniz and the hills of the castle and Graça neighborhood that was the area designated for the defeated Moors after the occupation of the city by D. Afonso Henriques as part of the Christian conquest over the Muslims in 1147. Due to its broken topography the neighborhood developed a layout irregular, the initial occupation surrounded the castle hill going down the slope to the water valley that existed in the past and is currently Almirante Reis Avenue. Mouraria is notable for its medieval layout, a reflection of Arab architectural influences, characterized by small streets, alleys, and narrow staircases. Additionally, the neighborhood's historical segregation, as it was not included within the Fernandina's wall in the 14th century, meant that it did not appeal to the wealthier classes of that period, (Da Costa, 2008).

Mouraria's historical activities were centered around its large Mosque, which was located where the Church of Nossa Senhora do Socorro is currently located. Additionally, in the area that is now Martim Moniz Square, there was once a smaller mosque along with public baths. This area held significant commercial importance, as Moorish craftsman served a large Christian customer. However, in 1496, the expulsion of Jewish and Muslim minorities from Portugal led to the transformation of Mosques and Islamic structures into churches and convents. These religious buildings were primarily utilized by the nobility and clergy, who played a pivotal role in incorporating Mouraria into the broader cityscape, (Langens, 2022).

Early in the 19th century, the neighborhood fell into poverty and crime, with bars and prostitution abounding. However, this tragic period was the perfect setting for the appearance of the traditional Portuguese musical genre called Fado, with which the fadista María Severa gave it better cultural state to the "Moro" neighborhood. Unfortunately, in the 20th century and in the name of the modernization of the area, the structure of the neighborhood was altered, collapsing several of the representative monuments of the 17th century, (Franco, 2016). Eventually, as part of Portuguese decolonization processes and current immigration phenomena, Mouraria is considered the multicultural enclave of the city center, being a commercial, gastronomic, and social center that allows multiethnic relations, (Oliveira, 2019). It is important to note, in recent years, the municipal policies have established strategies aimed at the regeneration of Mouraria. These strategies involve transforming its stigmatized image into that of a rehabilitated space, both in terms of infrastructure and by providing access to cultural activities. These initiatives are designed to enhance the neighborhood's identity, tradition, and diversity (Estevens et al., 2019).

2.4.4 Bairro Alto

The Bairro Alto throve between the end of the 15th and 16th centuries as a response to the revocation of the mandate that limited the city to develop only within the medieval Fernandina wall, which sought to make the central area of the city and the port the social and commercial axes. The modification allowed this neighborhood to begin the process of Lisbon as a modern city where the popular classes would begin their location externally, mostly on the riverbank in fish trades, which changed over time on the northernmost part. Palaces, convents, and churches led by merchants and bourgeois families of that time were located on this hill, (Da Costa, 2008).

With its evolution, Bairro Alto was characterized by being a neighborhood with greater ventilation than the city center, especially in times of the plague, and due to its topographic conditions, the rainwater flowed directly into the Tejo River due to its long and regular streets, which guaranteed better health conditions, however. with narrow streets that at the time reduced the conditions of characteristic mobility that even today persists as a neighborhood with a large commercial concentration, highlighting meeting areas and social recreation for the city's residents, (Nofre, 2020).

In contemporary times, Bairro Alto has earned a reputation as a bohemian enclave, attributed to the many journalists and literary authors who have resided there. However, starting in the 1990s, the neighborhood began to develop a vibrant nightlife, which has since become the city's premier nightlife spot. This evolution is largely due to the flow of students drawn by affordable rents and the convenience of local amenities (Baptista et al., 2018).

Chapter 3

3. Methodology

The methodology of this study centers on exploring the spatial definitions of three historic neighborhoods in Lisbon – Alfama, Mouraria, and Bairro Alto. This exploration is achieved through the analysis of data obtained from cognitive mapping, extracted via an online web-based survey, and geotagged online activity from user-generated content (UGC). Both datasets are integral components of the CityMe project, an initiative funded by Portugal's Fundação para a Ciência e a Tecnologia (FCT). The CityMe project is designed to deepen our understanding of how citizens mentally map the city, thereby enhancing spatial analysis and community participation in public policies and urban planning.

A key contribution of the CityMe project is the development of a fully functional public website (https://cityme.novaims.unl.pt/) that facilitates participant data collection through a web-based map survey and provides access to sources of geotagged content from UGC. The workflow of the study is divided into two main lines.

The first phase focuses on extracting representative boundaries of the three historic neighborhoods. This process involves collecting participant sketches through the web-based map survey. This developmental line details data collection and processing methods, extraction techniques for perceived boundaries, and culminates in the representation of the agreement on perceived boundaries of Alfama, Mouraria, and Bairro Alto.

The second phase of the workflow addresses the retrieval of representative boundaries from various sources of geo-tagged online activity. This part of the study outlines the specifics of data collection, content selection criteria, coding processes, clustering extraction, and ultimately, the generation of bounding shapes for the historic neighborhoods.

The final phase of the study is where these two phases converge, enabling a quantitative comparison of the spatial definitions derived from both cognitive mapping and geotagged online activity. This concluding phase involves calculating metrics to understand the relationship between the neighborhood extensions generated by the two methodologies.

The figure 2 illustrates the structure of the study methodology.

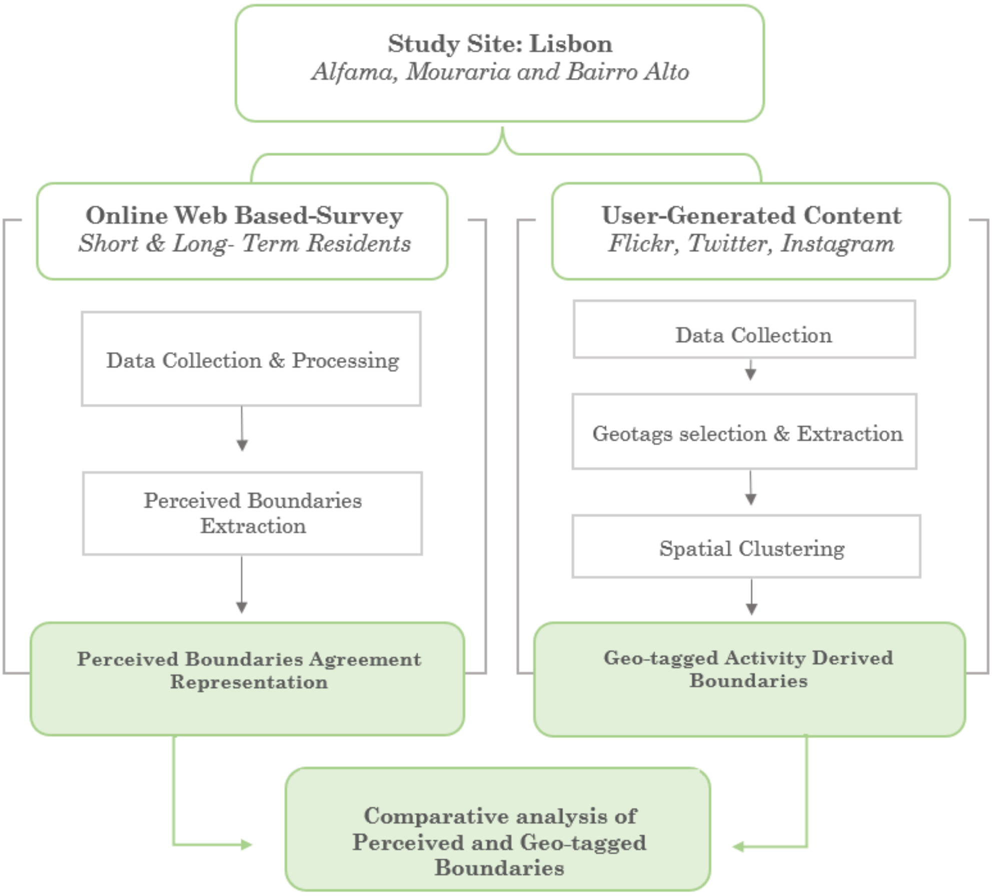

Figure 2. Schema thesis workflow

3.1 Online Web-Based Survey

3.1. 1 Data collection & Processing

The first phase of the project uses the results of the online web-based survey to extract the residents' cognitive perception of their spatial environments, and which represents vernacular neighborhoods in the city of Lisbon.

The questionnaire was shared on the project's twitter account *(https://twitter.com/CityMe5)* and made known to the student community, teachers, university team, as well as external professional connections for their participation; therefore, the contributions received are part of the dataset that has been selected to develop this *r*esearch.

The survey (https://cityme.novaims.unl.pt/survey) is compose of two sections: first, the demographic questionnaire that is related with respondent demographic variables and second, based on their connection with the city, asking participants to draw areas corresponding to the citizens mental representation about geographic extensions of each historic neighborhood on the base map of Lisbon.

In the context of this project, aimed at achieving established objectives, two key variables were selected from the survey. The first variable relates to the length of residence in the city, captured by the question 'How long have you lived in Lisbon?', which presents a series of time intervals as shown in figure 3. This variable is crucial for the project's development, enabling the evaluation of differences between two categories for further analysis: short-term and long-term residents of the city of Lisbon. A ten-year threshold serves to distinguish between short-term and long-term residents based on their selections from the survey's time interval options. Short-term residents are identified as those respondents who chose options indicating they have lived in Lisbon for up to 10 years. Conversely, long-term residents are defined as those selecting options showing they have resided in the city for more than 10 years.

Figure 3. The 'About you' section of the CityMe survey, related to options for the length of residence in the city of Lisbon (Source CityMe, 2021).

It should be noted that there is a degree of arbitrariness in the definition of the response categories of the survey from which the data was collected; however, Earlier investigations have scrutinized the strong connection and identification people feel with their locations based on how long they have lived there. For example, research cited in this study (Hernández et al., 2007) demonstrates that long-term residents in specific neighborhoods exhibit stronger place attachment than place identity. While both attachment and identity to a place vary, factors like duration of residency, birthplace, and the context of the survey (such as the neighborhood or city) influence these variations. Additionally, (Tang et al., 2021) study highlights the spatial aspect, noting it is influenced by social interactions over time. These interactions define the boundaries of a place and the sense of belonging among individuals. The study also reveals that the meanings attributed to a place distinctly vary with the length of residency.

The second aspect involves the geometries captured in the survey, which are based on residents' perceptions of the locations of the historical neighborhoods to be analyzed: Alfama, Mouraria, and Bairro Alto, as illustrated in Figure 4.

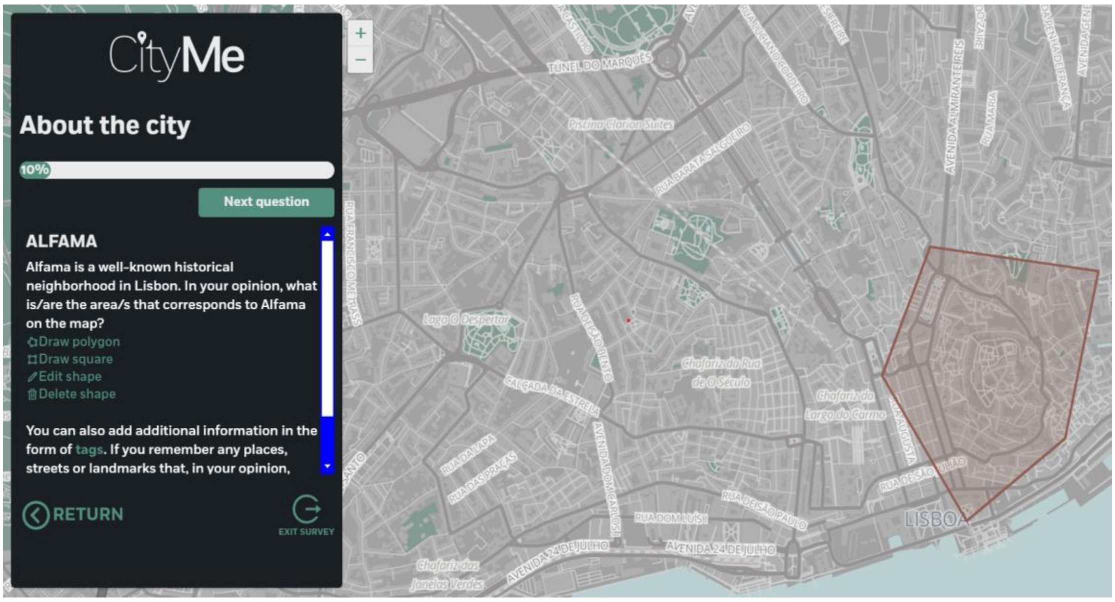

Figure 4. The 'About the city' section of the CityMe survey, related to the capture of the extent of historic neighborhoods such as Alfama (Source: CityMe, 2021).

The processing data began accessing to the survey responses database to know its structure then proceeding to extract the answers associated with the question about the length of residence and geometry's location opinions previously mentioned using SQL Query according to the database structure which have an identification code for consultations (https://github.com/CityMe-project/CityMe-survey_data), This process is illustrated in Figure 5.

```
'QUESTION 2.1.1 REPRESENTS ALFAMA NEIGHBORHOOD'

select * from main.v_all_geometries WHERE question = '2.1.1'
and survey_id in(select survey_id from main.v_all_answers
WHERE "1.2.2" = 'between 10 and 20 years'OR "1.2.2"='more than 20 years')
```

Figure 5. Example of selection of geometries for long-term residents in Alfama neighborhood using SQL Query.

The datasets retrieved, containing geometries perceived as historical neighborhoods, were as follows, Table 1 .

Neighborhood/ Resident groups	Short-term residents	Long-term residents
Alfama	56	80
Mouraria	31	66
Bairro Alto	32	63

Table 1. Number of geometries extracted from the survey by neighborhood and resident groups.

During the data processing phase, records with null fields, those outside the study area, and records considered outliers in terms of the average polygons of the dataset were eliminated. This refinement led to the final datasets, distinguished by the length of residence in Lisbon. Short-term residents, defined as individuals who have lived in Lisbon for less than 10 years, contributed 55 responses from Alfama, 29 from Mouraria, and 32 from Bairro Alto. In contrast, long-term residents, those who have resided in the city for more than 10 years, submitted a larger volume of responses, with 74 from Alfama, 62 from Mouraria, and 63 from Bairro Alto. Figures 6 , 7 and 8 illustrate the spatial distribution of these final datasets, showcasing the contributions from both resident groups in every historic neighborhood.

Alfama

Figure 6. The spatial distribution of Alfama's sketches.

Mouraria

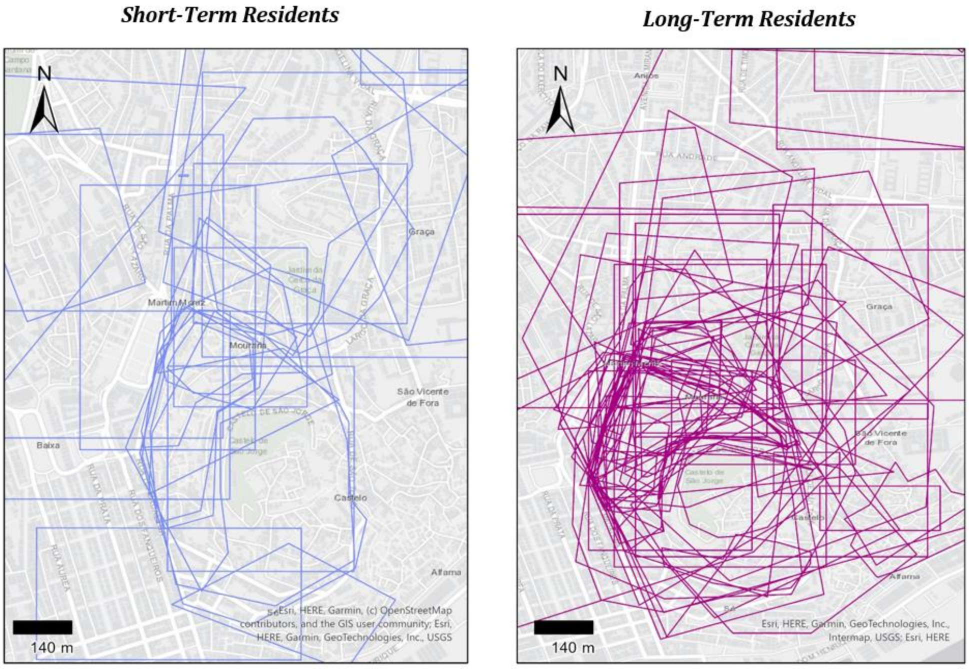

Figure 7. The spatial distribution of Mouraria's sketches

Bairro Alto

Figure 8. The spatial distribution of Bairro Alto sketches.

3.1. 2 Perceived Boundaries Extraction

To accurately represent the perceived boundaries of each neighborhood based on respondent agreement, we quantified the number of intersecting surfaces by counting the overlapping areas in input sketches through the overlay analysis adopted by Brown in his study named Mapping and measuring place attachment (Brown et al., 2015) . This approach allows us to establish consensus thresholds and effectively delineate neighborhood limits as perceived collectively by respondents.

For the extraction of perceived boundaries, the use of GIS tools was essential, which allowed the definition of consensus areas by quantifying the overlap between the different drawn polygons. The process began inputting each of the datasets and applying the following steps:

First, the Union tool was applied, creating new polygons for each unique area of overlap. This means that when two polygons overlap, a new polygon is generated to represent just their overlapping area, and this process is repeated for any additional polygons overlapping in the same area. Subsequently, some of the resulting polygons from this operation might have been multipart features, meaning a single record in the attribute table represents multiple disconnected geometries. The Multipart to Single Part tool was then applied to break these multipart features into individual features, with each part becoming a separate feature in the output layer. This step was important; since, it ensures that each distinct polygon, including those created from overlaps, is treated as an individual entity. Finally, a Spatial Join was performed on the previous output, creating a Join Count column. This column, generated from the process, indicates the number of polygons that overlap at each location.

The agreement within the extracted polygons is represented by the *Join Count value*. A higher Join Count value implies a greater degree of agreement or overlap among the input polygons in that area. This can be interpreted as a stronger consensus or a more commonly recognized boundary in the context of neighborhood delineation and spatial relations. The overlay analysis steps were applied to each feature class, which represented each dimension for both participant groups, short-term and long-term. Figure 9 shows the overlap degree between the already processed polygons according to class classification by color graduation.

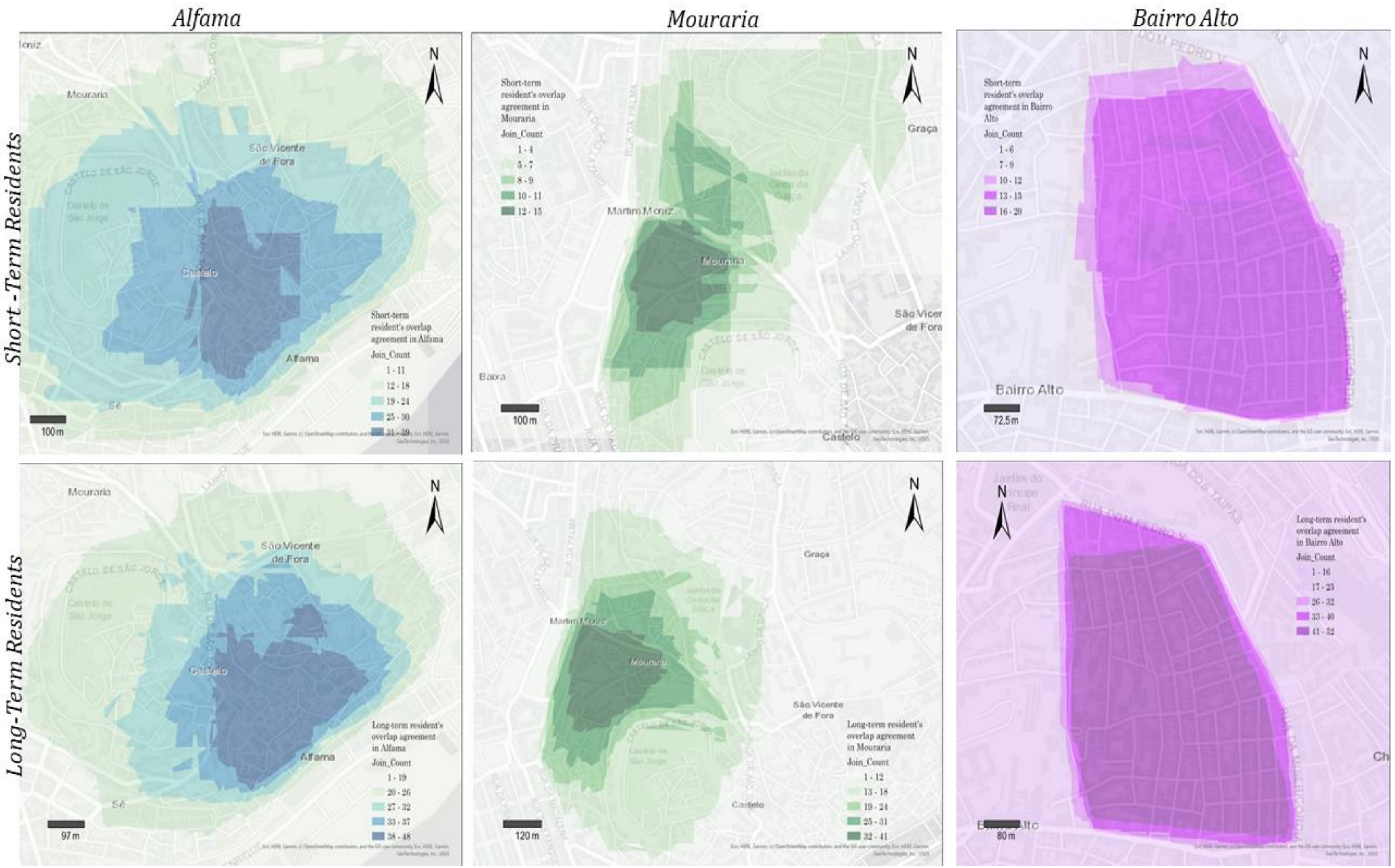

Figure 9. Overlap agreement maps by polygon count for short- and long-term residentsin Alfama, Mouraria, and Bairro Alto historic neighborhoods!

This study employed the Minimum-Maximum standardization method to make uniform the count field of overlapping entities, establishing values ranging from 0 to 100 (Percentage). Consensus thresholds were chosen, supported by the literature, which shows percentages that embody the resident's delineation boundaries and their identity effects using sketch maps (Bae & Montello, 2018).Conforming to the study, the areas of greatest agreement were >75 percent which we will call the ***Core*** region, and areas indicating at least 50 percent but no more than 75 percent agreement, it will call ***Domain*** region.
The categorization into "Core" and "Domain" regions was chosen based on Meinig's contribution, who the extent of the region occupied by Mormon culture over time in concentric circles, focusing on the state of Utah and southeastern Idaho in the United States of America. His model, as presented in 1965, is adopted and adapted to the neighborhood level in this study, (Meinig,1965). The integration of these categories with consensus thresholds is depicted in Figure 10.

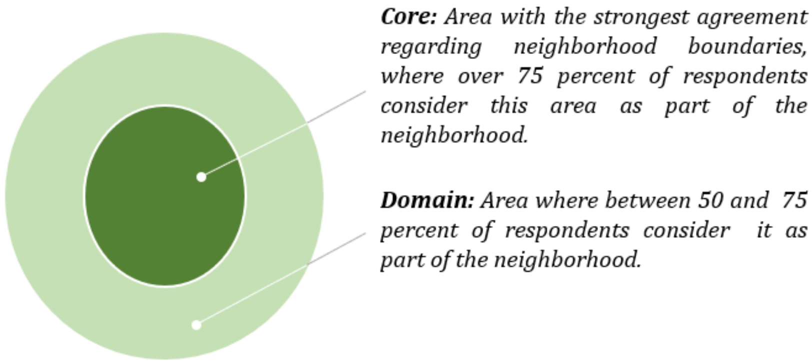

Figure 10. Neighborhood Regions Scheme

At this stage, the outcome yields the study's ground truth, which becomes an invaluable primary source encapsulating unified boundaries and residents' detailed perceptions of neighborhoods. This ground truth serves as a benchmark for evaluating the accuracy of maps produced from user-generated content, and it plays a vital role in discerning the differences and similarities between local perceptions and digital depictions developed in the next stage of the study.

3.2 User Generated Content

3.2.1 Data collection

The study incorporates a second dataset derived from user-generated content. This dataset is composed by geotagged posts, pages, advertisements, points of interest (POIs), and their respective attributes, including textual data, all of which were gathered from a variety of social media platforms, it's important to note that the data extracted from these sources came from public profiles of users. The extraction of this data was made possible using both official and unofficial APIs, and it was undertaken as an integral part of the CityMe project.

The CityMe project offers a unique and comprehensive perspective on the spatial dynamics of Lisbon by combining surveyed data (Using in the first stage of this research) with user-generated content from social media platforms. This approach provides a more holistic understanding of the area; in consequence, A key element of the CityMe project UGC section is the City Me-UGC online repository, accessible at *https://github.com/CityMe-project/CityMe-UGC,* enabling to perform spatial analyses that compare and characterize regions based on both surveyed data and the user-generated content.

In this segment of the study, the focus is on the methodologies used for extracting representative boundaries from user-generated content. Data was gathered from three social networks Flickr, Twitter, and Instagram specifically targeting mentions of the names of historic neighborhoods relevant to this study within the textual content. The extraction process was carried out through Python coding. This approach allowed for a comprehensive analysis of user interactions and references to these historic neighborhoods.

The following content provides a brief overview of the structure of the three sources and describes the unique characteristics and contributions of each platform to the research.

Flickr

Flickr operates as a digital platform for photo management, renowned for its extensive collection of geotagged images, as noted by (Gao et al., 2017a). On Flickr, users have the capability to upload and disseminate their photographs, either within specific Flickr groups or to a broader audience. These images are often accompanied by user-generated tags, providing descriptive context, and are geotagged with precise coordinates. This feature is particularly useful for research, as Flickr is predominantly utilized by tourists, which aligns with Lisbon's status as a sought-after travel destination. The georeferencing of photographs on Flickr can be accomplished either automatically or manually, leveraging the location-based technologies embedded in smartphones and cameras.

Twitter

Twitter is recognized as a prominent social network platform, primarily utilized for sharing opinions and news reporting. According to (Li et al., 2013), user-generated data from Twitter offers valuable insights into the characteristics of various locations and the demographics of people who frequent these places, whether they are residents, workers, or tourists. This aspect of Twitter makes it a significant source for understanding social dynamics and public perception.

Furthermore, as pointed out by (Salvatore et al., 2021), Twitter data can serve as a supplementary source of information, complementing statistical data obtained from official sources. The brief, yet expressive messages on Twitter often reflect people's instant reactions and perceptions about various aspects of their environment, such as neighborhood safety and crime levels. These tweets not only provide individual perspectives but can also trigger group reactions, offering a real-time glimpse into public sentiment.

Instagram

The shared texts on Instagram offer several benefits, as outlined by (Bernabeu-Bautista et al., 2023). These include insights into the temporal and spatial trends of people at specific locations, the spontaneous reactions and perceptions individuals have of their surroundings now of posting, and the identification of patterns related to activities in urban spaces. Such patterns reveal the diversity of preferences, mobility, and identity of cities at various points of interest. Initially, the selection of georeferenced content on Instagram was done manually. However, this process later evolved to utilize integrated positioning technology, enhancing the accuracy and efficiency of georeferencing. This technological shift has allowed for a more precise and comprehensive understanding of how people interact with and perceive their urban environments.

3.2.2 Geotags selection & Extraction.

In this research, we focused on standardizing the data attributes gathered from interactions on Flickr, Twitter, and Instagram. The key variables standardized were the mention of neighborhood names within the textual content and their geographical locations. This data was compiled over a decade, from 2012 to 2022, a period during which content related to the city of Lisbon was available across all these platforms. The unfiltered datasets comprised: 5,669,531 records for Twitter, 50,946 for Flickr, and 35,657 for Instagram. The standardization process was essential for ensuring consistency and accuracy in the analysis of the data collected from these diverse social media sources.

To obtain the relevant neighborhood description of geotagged data from social platforms, names of neighborhoods as keywords were extracted from the textual attributes of the geotagged data such as the "status", "description" and "tags" by filtering out exact mentions of the keywords and associated extracted neighborhood keywords (Alfama, Mouraria and Bairro Alto) with the point coordinates. In case of multiple neighborhood names were mentioned in the same data entry, Euclidean distances between the data point and centroids of mentioned neighborhoods (Using ground truth already generated as reference) were calculated and the nearest neighborhood to the point was assigned as the neighborhood attribute of the point.

Coordinates of points from certain platforms such as Instagram and Twitter are not organic, where geotagged coordinates of the data are not retrieved from the user's actual positioning system but rather assigned with a pre-populated coordinate associated with the name of the place the user is checking in. This nature assigns redundant and overlapping coordinate points to the geotagged data regardless of the user's actual location. This becomes problematic when calculating the (minimum) distance band between neighboring points in identifying and creating clusters since a significant amount of points have identical coordinates which affects the statistical calculations resulting near sub-meter distances for the distance band. These identical/redundant/overlapping points were removed till only one unique point per coordinate is left.

To address the imbalance in geotagged data from sources like Twitter, which significantly exceeds that from other sources, our study implements a method to calibrate the data volume. Following the approach by (Gao et al., 2017a), which identified a similar discrepancy, we first establish the (average/maximum) contribution level of normal or ordinary users. This level is determined by calculating the data or posts produced by the top 10% of active users, representing the 90th percentile of the user base. A threshold is then set at this level, and any user contributions exceeding it are scaled down accordingly. Utilizing this adjusted threshold ensures that each user's data contribution is proportionate and consistent within the dataset, which is then used for further analysis.

As a result, we obtained the filtered datasets with the three names of historical neighborhoods from every source. Important to mention that the Flickr registers calling for Mouraria did not yield results. Looking for an explanation, it was found that Mouraria was mentioned, but the location of the post was caught by Alfama centroid influence area. Figure 11 illustrates the quantities derived after performing the data mining process. Figure 12 displays the extracted data and a representation of the postings used as the extraction source.

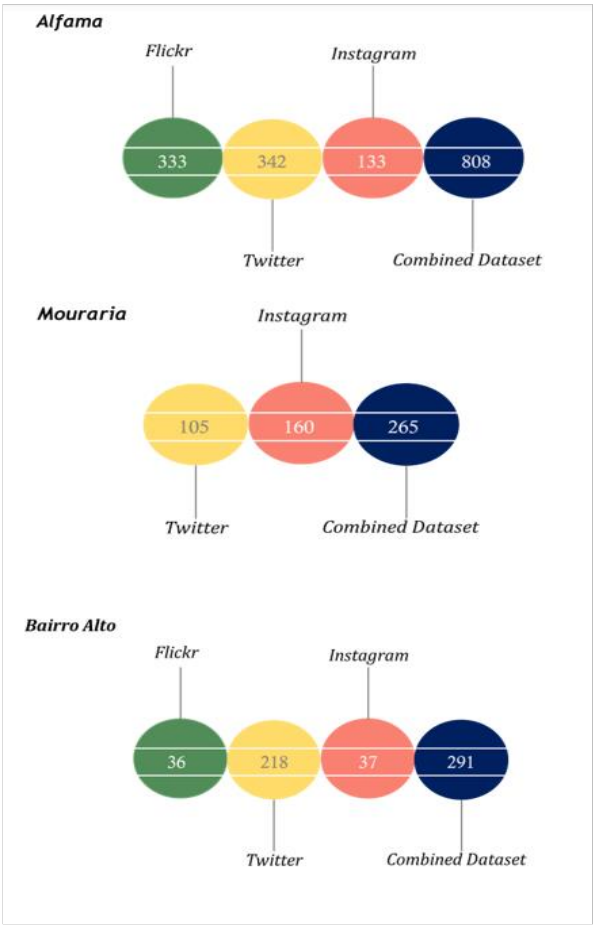

Figure 11.Number of posts extracted based on the name of each neighbourhood from every UGC source.

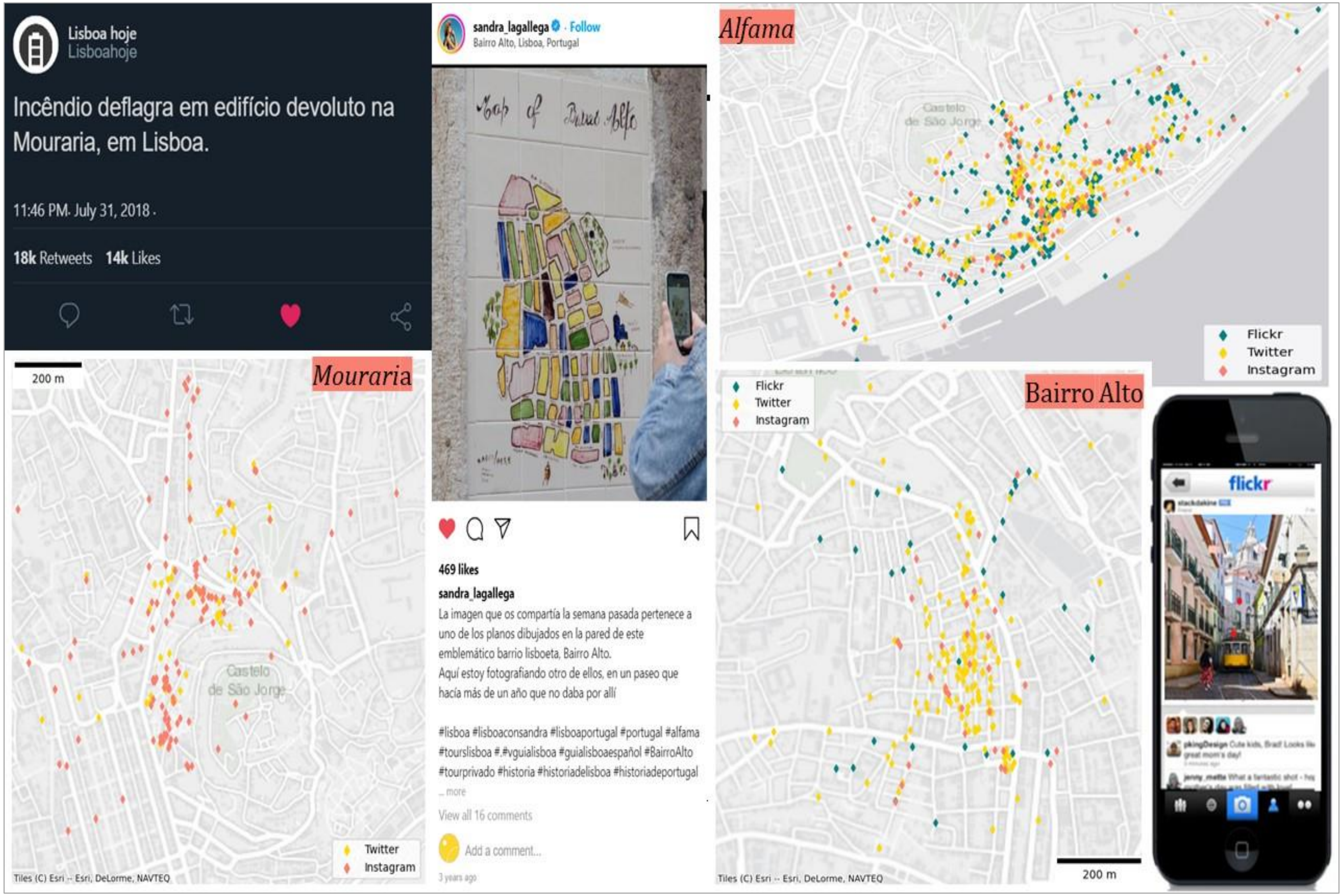

Figure 12. Distribution and depiction of selected posts from three different UGC sources across the historic neighborhoods of Lisbon.

3.2.3 Spatial Clustering

By extracting a dataset of geo-referenced social media postings from three different sources, each categorized by neighborhood, the research concentrated on employing a coding process for identifying clusters of points. This goal was accomplished using the Approximate DBSCAN (A-DBSCAN) clustering algorithm, developed by (Arribas-Bel et al., 2021)
.The study aimed to gain insights into urban areas by analyzing the density of points, particularly those associated with buildings within these areas. For the spatial clustering in this research, pre-extracted geotags were utilized, facilitating the identification of high-density clusters from each source, and defining their boundaries.

A-DBSCAN algorithm is useful to the calculation and generation of shapes that represent the spatial extent of neighborhoods, illustrating the clustering patterns of social media geotags, as highlighted by (Tang et al., 2022). The application of this algorithm requires two input parameters. The first one, known as the ***eps*** parameter, represents the maximum distance between two points necessary for them to be considered part of the same neighborhood or cluster. Firstly, a nearest neighbor analysis was performed on the dataset. This analysis involves calculating the distance from each point in the dataset to its closest neighbor. The purpose of this step is to understand the typical distances between points in the dataset, which is crucial for determining the appropriate scale for clustering.

After completing the nearest neighbor analysis, the next step was to calculate a specific percentile of these distances as in (Tang & Painho, 2023, Tang et al., 2022) studies. The calculated percentile value is then used as the 'eps' parameter in the clustering algorithm. The choice of which percentile to use can depend on the dataset and the desired sensitivity of the clustering algorithm to outliers or noise. That was the case for the current study were the percentile values oscillate between the 90^{th} to the 99^{th} percentile per social platforms.

The second parameter is called ***'Minimum Points'*** and refers to the minimum number of neighboring points required for a point to be considered part of a cluster. Given that each data source exhibits variations in the number of entries, we opted to use percentages for choose Minimum Points as in (Gao et al., 2017) research. In this study, the percentages chosen were 3% and 5% of the total number of postings per data source, to model the vague nature of cognitive regions.

Finally, the A-DBSCAN algorithm was run iteratively. Table 2 shows the combination of parameters by source and neighborhood to generate a stable and representative delineation based on the distribution of geotags. The construction of that polygons was then undertaken to approximate the bounding shape of the historic neighborhoods, using the **α-shape** algorithm due to its ability to provide a smoothed delineation,(Edelsbrunner et al., 1983). It is important to note that in addition to calculating polygons for each platform it was created a shape derived from ***all sources*** for each of the historic neighborhoods under study.
The project aims to reveal patterns of consistency and variance in the identification of neighborhood boundaries by comparing this primary data with the insights derived from user-generated content analysis, as discussed in this phase. These findings will be further explored in the subsequent phase.

	Flickr		**Twitter**		**Instagram**		**Combined**	
NBHD/Parameters	**eps (%)**	**mins(%)**	**eps (%)**	**mins(%)**	**eps (%)**	**mins(%)**	**eps (%)**	**mins(%)**
Alfama	0.99	0.03	0.99	0.03	0.99	0.03	0.99	0.05
Mouraria			0.95	0.03	0.90	0.05	0.95	0.05
Bairro_Alto	0.99	0.03	0.90	0.03	0.90	0.03	0.90	0.05

Table 2. A-DBSCAN parameters values by source and neighborhood.

3.3 Comparative Analysis of Perceived and Geo-tagged Boundaries

The final stage of the study integrates the initial segments, undertaking a quantitative comparison of spatial definitions derived from cognitive mapping and geotagged online activities within the historic neighborhoods of Alfama, Mouraria, and Bairro Alto. This comparison aims to measure the agreement between neighborhood extensions as identified by each method. Central to this analysis are two metrics: the Intersection Over Union (IOU) and the F-scores.

The IOU metric quantifies the overlap between two shapes, with values ranging from 0 to 1 where 0 represents no overlap and 1 indicates perfect overlap. Similarly, F-scores assess the accuracy and relevance of this overlap by calculating precision and recall, providing a balanced measure of a methodology's effectiveness in accurately capturing and comprehensively covering neighborhood boundaries as was calculated in (Tang et al., 2022). F-scores also range from 0 to 1, where 0 signifies the worst precision and recall, and 1 the best.

Calculations for both IOU and F-scores are grounded in the "ground truth" established by perceived boundaries obtained in the first stage of the study and the polygons generated from user activities. Python was utilized for these computations, facilitating the coding and analysis of the results. This approach provides a quantitative basis for discussing the spatial definition of historic neighborhoods, allowing for a comparison of the insights gained from both cognitive mapping and geo-tagged data.

Chapter 4

4. Results & Discussion

4.1 Perceived Boundaries Agreement Representation

From the surveys, the extraction of perceived boundaries successfully identified the core and domain regions of the three historical neighborhoods. The ensuing maps, enriched with crucial insights regarding their extents, are set to highlight the variations as perceived by different resident groups. The main findings from the extracted delineations are presented below.

4.1.1 Alfama

The map displays the perceived boundaries of Alfama as delineated by short-term and long-term residents, figure 13 and Table 3 depicts its extension in Km^2. Short-term residents have a broader, less defined perception of the neighborhood's boundaries. This expansive domain interestingly includes the Castelo de São Jorge and Bairro de Santa Cruz described in (Franco, 2016) route, indicating an inclusive approach to defining neighborhood limits. In contrast, the map for long-term residents shows a more compact domain, excluding Castelo de São Jorge, which could reflect a more nuanced understanding of Alfama's traditional limits described by (Da Costa, 2008) study.
Both groups identify Alfama's core similarly, though the core for long-term residents aligns more closely with the medieval street layout and its location in an area with a complex topography,(Tang & Painho, 2023). This indicates a possibly deeper connection to and understanding of the neighborhood's historical and functional core, highlighted by points of interest such as the Church of São Miguel and Santo Estêvão, Rua da Regueira at the center of the core region, and extending up to Rua Escolas Gerais .(The complete location of Alfama's points of interest can be found in the Appendix).

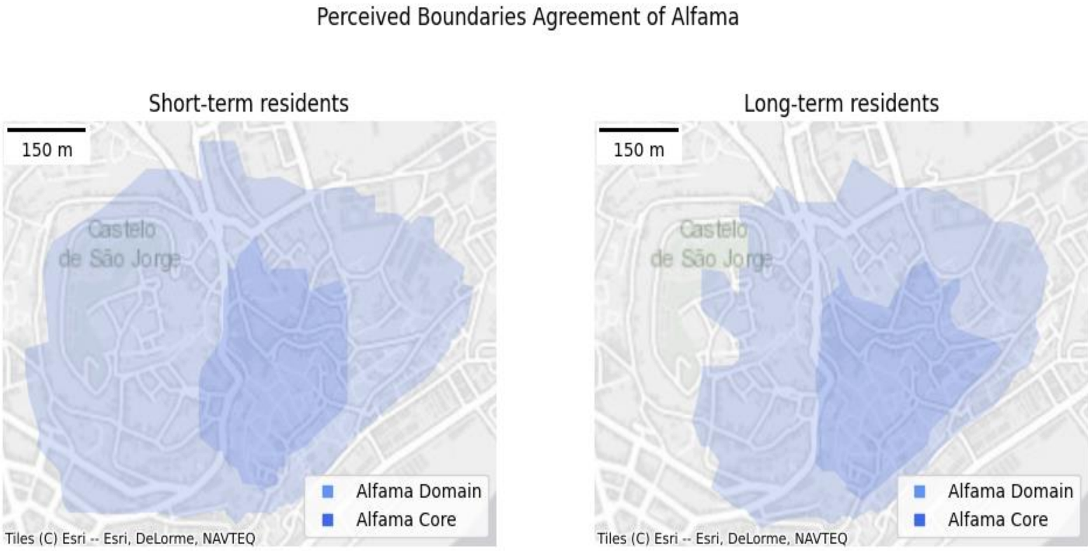

Figure 13. Map of Alfama showing perceived boundary agreement between short-term and long-term residents, highlighting consensus regions.

Region	Short-Term Residents Km^2	Long-Term Residents Km^2
Domain	0.371	0.280
Core	0.079	0.082

Table 3. Area of consensus regions by resident groups in Alfama.

4.1.2 Mouraria

Figure 14 showcases the perceived boundaries of Mouraria as defined by both short-term and long-term residents Table 4 depicts its extension in Km^2, revealing a slight difference in the domain areas, with short-term residents' domain measuring 0.1251 square kilometers and long-term residents' at 0.127 square kilometers. This indicates a broadly similar understanding of Mouraria's extent among both groups, albeit with long-term residents considering a marginally larger area.
A more pronounced difference emerges in the core regions, where long-term residents have identified a larger core area of 0.050 square kilometers, as opposed to the 0.033 square kilometers by short-term residents, suggesting long-term residents have a more expansive view of Mouraria's central part.
However, the delineation of both the domain and core regions differs significantly. Mouraria, evolving over time, exhibits more diffuse boundaries due to its rapid transformation influenced by a diverse social and economic composition, as (Oliveira, 2019) notes. A notable commonality in the core area for both resident groups is Rua Marquês de Ponte de Lima, recognized for hosting the great mosque, now only evidenced by archaeological remains,

according to (Langens, 2022). This historical element underscores the blend of tradition and change that characterizes Mouraria's urban landscape.

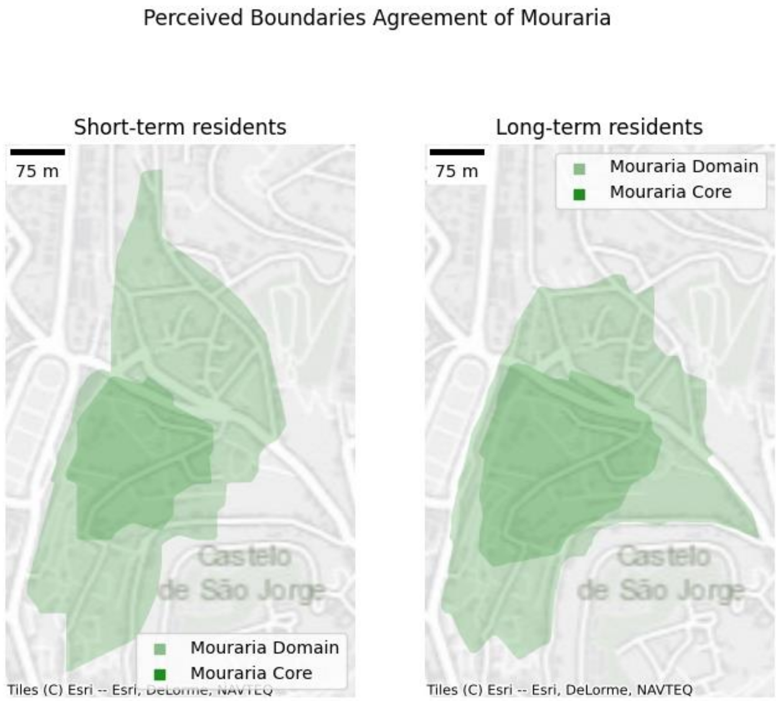

Figure 14. Mouraria map showing perceived boundary agreement between short-term and long-term residents, highlighting consensus regions.

Regions	Short-Term Residents (Km^2)	Long-Term Residents (Km^2)
Domain	0.1251	0.127
Core	0.033	0.050

Table 4. Area of consensus regions by resident groups in Mouraria.

4.1.3 Bairro Alto

The maps comparing the perceptions of Bairro Alto by short-term and long-term residents reveal key differences in how each group views the neighborhood's boundaries as figure 15 depicts, Table 5 shows its extension in Km^2. Long-term residents have a broader understanding of both the core and the domain areas, with the domain being 0.189 km^2 and the core 0.153 km^2, which are larger than the areas perceived by short-term residents (0.178 km^2 for the domain and 0.095 km^2 for the core). The consistent shapes observed in the maps of both resident groups specially the domain region expose a shared general understanding of the neighborhood's layout. The aforementioned allow us to trace back to the origins of the neighborhood, which emerged from the construction of residential homes for the nobility, exhibiting a structure that deviates from the medieval layout, (Nofre, 2020). This is corroborated by a study conducted by (Tang & Painho, 2023) on the same area, which reveals that the distribution of the street network significantly influences the neighborhood's legibility for both recent and long-standing city residents.

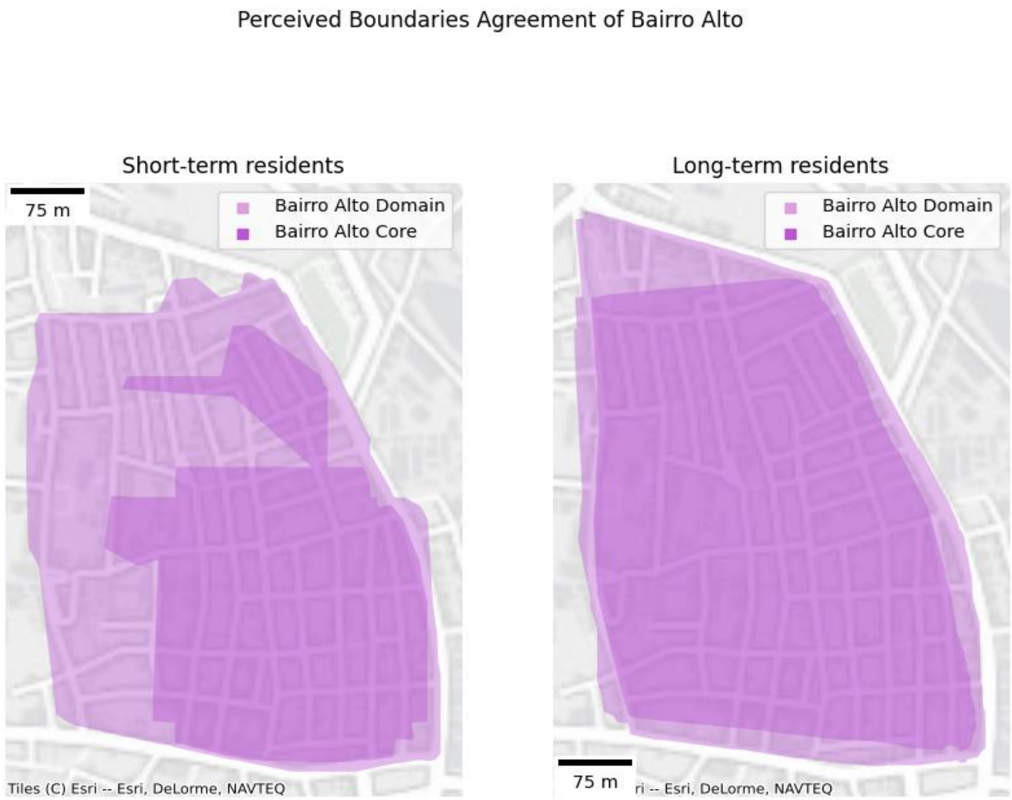

Figure 15. Bairro Alto map showing perceived boundary agreement between short-term and long-term residents, highlighting consensus regions.

Regions	Short-Term Residents (Km^2)	Long-Term Residents (Km^2)
Domain	0.178	0.189
Core	0.095	0.153

Tabla 5. Area of consensus regions by resident groups in Bairro Alto.

4.2 Geo-tagged activity derived Boundaries.

After completing the clustering process and configuring the shapes for all datasets across the three historic neighborhoods in the second stage of the study, the resulting maps were generated. These maps display colored shapes derived from this data to visually represent the varying concentrations and patterns of activity within each neighborhood.

4.2.1 Alfama

The frequency of mentions of Alfama across various social media networks reveals a very broad extension in the shape outputs in that area of the city (figure 16), propose a vagueness in the precise notion of the neighborhood's location. This phenomenon is attributed to Alfama's reputation as one of the main tourist areas in the city, leading to its name being referenced even in places that are not Alfama, as (Cocola-Gant & Gago, 2021) indicate. All sources follow a similar pattern of distribution, skirting around the hill of Castelo de São Jorge and extending west to the city center. The surface area generated from the aggregation of all sources is more consistent with the core of Alfama, and notably, its left edge coincides with several renowned tourist points of interest, such as Rua do São Tomé and the Miradouro de Santa Luzia highly visited for preserving part of the medieval wall, (Franco, 2016).
(The complete location details of Alfama's points of interest can be found in the Appendix)

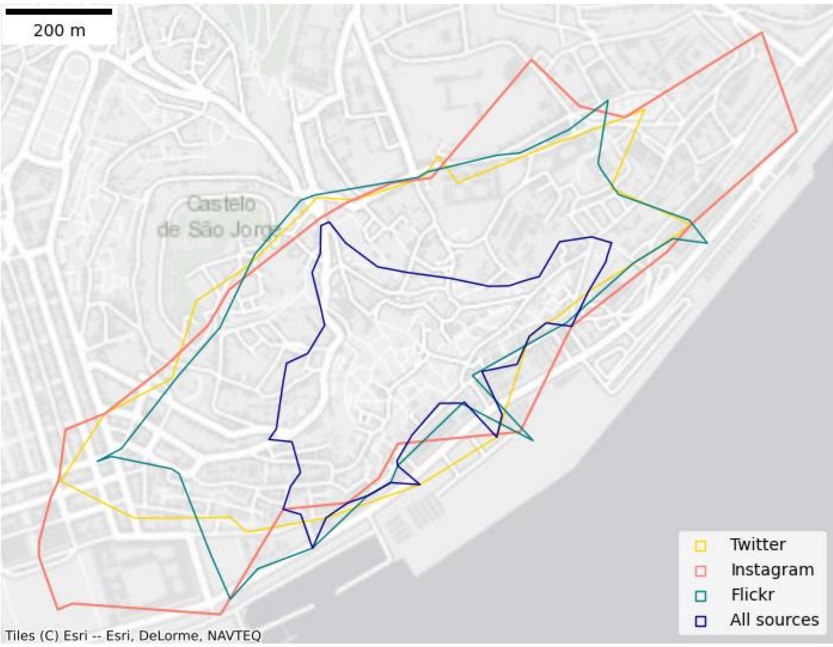

Figure 16. Geo-tagged activity derived boundaries of Alfama.

4.2.2 Mouraria

The results of the surfaces representing the spatial agreements of Mouraria as figure 17 depicts vary significantly across different social media networks. This variance is rooted in the clustering process used to obtain a stable and representative surface, which proved to be challenging. The initial distribution of mentions was sparse or even non-existent, as was the case with the social media network Flickr. For instance, the Instagram output forms a wide shape, which is the result of a very dispersed source of data for its conformation in Mouraria. When comparing the three outlines, it becomes evident that each social media platform exhibits a distinct pattern of user activity. These patterns are likely influenced by the type of content shared on each platform and Mourarias' unique internal social dynamics , as previously analyzed. However, the results are in places of historical character for the neighborhood, although not in a similar way for all platforms.

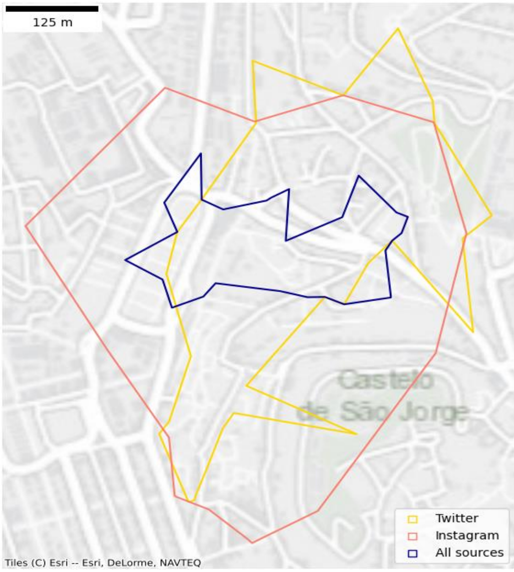

Figure 17. Geo-tagged activity derived boundaries of Mouraria.

4.2.3 Bairro Alto

The social media outputs from Instagram, Twitter, and aggregated sources in Bairro Alto present similar patterns in shape as figure 18 depicts, indicating a possible consensus among users of these platforms on the areas of interest within the neighborhood. Interestingly, Flickr's output displays a wider shape, reflecting the dispersion of user interactions. This showing a more inclusive approach in representing Bairro Alto, although it may not recognize specific areas as precisely as the other platforms.

On the other hand, the outline derived from Twitter data is the most contained and streamlined, closely following the inner parts of the neighborhood. Both Twitter and Instagram data suggest that user activity in Bairro Alto is centered around particular hotspots or venues. This is supported by the fact that popular nightlife areas, such as Rua da Rosa, Rua da Atalaia, Rua São Pedro de Alcantara, and traditional restaurants such as Tasca do Chico, are located within these outlines, as confirmed by (Nofre, 2020) .(The complete location of Bairro Alto's points of interest can be found in the Appendix).

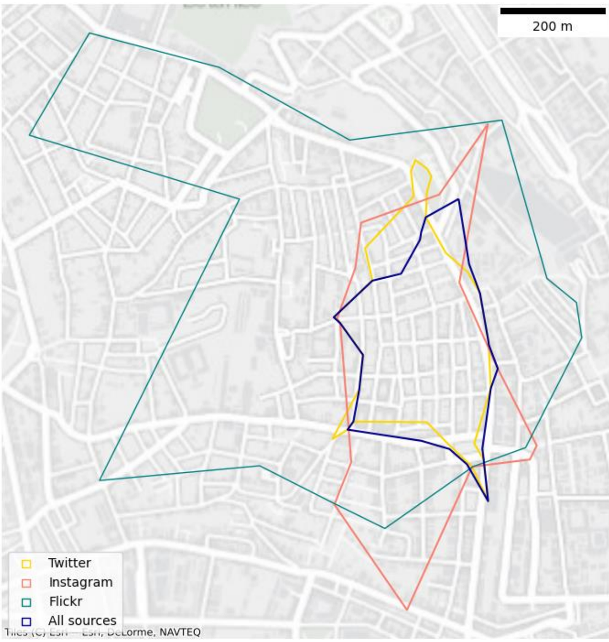

Figure 18. Geo-tagged activity derived boundaries of Bairro Alto.

4.3 Comparative Analysis of perceived and Geo-tagged Boundaries

The third phase encompasses a comparative analysis. This section outlines the computation of metrics, including the Intersection Over Union (IOU), which measures the extent of overlap between the identified shapes, as well as the evaluation of overlapping areas and F-scores for each dataset. Utilizing the perceived boundaries defined as the benchmark in the study's initial phase, these are compared against the user-generated shapes to determine the IOU and F-scores. The maps and metrics are accompanied by observations, highlighting the most significant findings by historic neighborhood according to residents' groups and regions of consensus.

4.3.1 Alfama

The comparative analysis of social media data in relation to the perceived boundaries of Alfama as defined by short-term and long-term residents reveals distinct patterns in both the domain and core regions. The results are display in the figures 19 and 20 and tables 6 and 7.

For short-term residents

In the domain region , Flickr stands out with the highest Intersection over Union (IOU) score of 0.538 among all sources, indicating a moderate overlap with residents' domain boundaries, see Table 5. Meanwhile, Twitter shows the highest recall value of 0.755, capturing a significant portion of the domain region but also including areas beyond it. When considering the combined data sources, the precision score peaks at 0.773, suggesting that the aggregate of all data most accurately aligns with the domain areas recognized by residents.
Moving to the core region, while all sources capture aspects of the core that residents agree with a perfect recall , they also encompass many areas residents do not recognize as part of it. Precision values are low across all platforms, with the combined data achieving the highest precision at around 40%, indicating that nearly half of the areas identified by all sources fall within the stricter consensus area. The F-score for the combined data in the core region reaches 0.549, which is the highest yet reflects the challenges in pinpoint accuracy for an area perceived narrowly.

For long-term residents

Flickr again depicts the highest IOU score of 0.580 in the domain region, showing a strong correlation with the areas that long-term residents identify with Alfama, see Table 6. The combined data has an IOU score of 0.394, slightly surpassing Instagram, and a precision of 0.766, the highest among all platforms, indicating a more accurate depiction of consensus areas when all data sources are considered. Flickr also exhibits a high recall of 0.912, capturing most of the domain region as perceived by residents. Furthermore, Flickr's precision F-score is significantly high at 0.734, capturing and comprehensively covering neighborhood boundaries for long-term residents. In the core region, IOU scores drop markedly across all platforms. Despite including all areas that long-term residents agree are part of Alfama, all platforms also capture areas outside of the consensus. The combined data's F-score is the most notable at 0.612, showing that the aggregation of all platforms, despite individual lower precision, offers a more balanced measure of accuracy between precision and recall for the core region.

Based on the metric results, Flickr's delineation of Alfama emerges as the most representative of the neighborhood, particularly when compared to the spatial agreement of long-term residents in the domain region. The data suggests that Flickr users, who are likely to be photography enthusiasts, concentrate their postings on very specific sites within Alfama. The behavior of users on photography-focused platforms like Flickr typically gravitates towards the most visually appealing locations within a neighborhood historical sites, viewpoints, and places of significant interest. This pattern aligns with the observations made by Hollenstein and Purves (2010), who also identified a similar trend in user interactions in their study. In their research, they attempted to use Flickr's place names and descriptions to delineate vernacular boundaries, as the methods employed in our study. Furthermore, Alfama's reputation as Lisbon's most picturesque neighborhood, as described by (Baptista et al., 2018), reinforces the idea that visually driven platforms like Flickr would naturally align closely with the areas that residents and visitors find most appealing.

The final delineation map of the historic Alfama neighborhood, based on Flickr geotags and the perceived boundaries by long-term residents, is detailed in the Appendix. As a support in the results descriptions this map showcases renowned points of interest, major road arteries, and images that capture the picturesque and historic character defining Alfama.

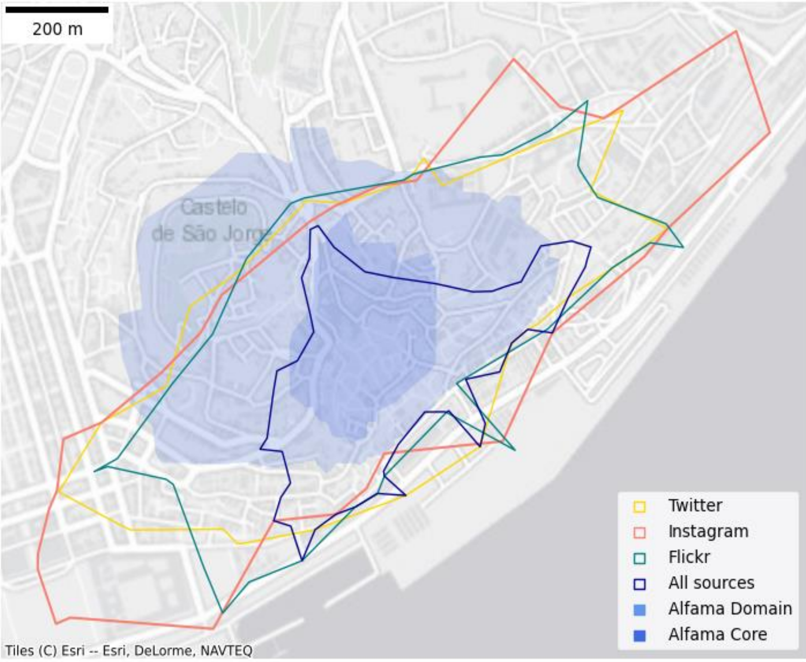

Figure 19. Comparative map of perceived and Geo-tagged boundaries for short-term residents in Alfama.

<u>Residents</u>

	<u>Domain</u>	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,538	0,523	0,385	0,310
	Overlap Km^2	0.276	0.281	0.274	0.127
	Recall	0,741	0,755	0,736	0,341
Short term	Precision	0,662	0,630	0,447	**0,773**
	F-score	0,699	0,687	0,556	0,473
	Core	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,191	0,179	0,130	0,378
	Overlap Km^2	79708,581	79708,581	79708,581	66878,3655
	Recall	1,000	1,000	1,000	0,839
	Precision	0,191	0,179	0,130	0,408
	F-score	0,321	0,303	0,230	0,549

Table 6. Quantitative results between user generated content datasets boundaries and Alfama Short-term residents ground truth.

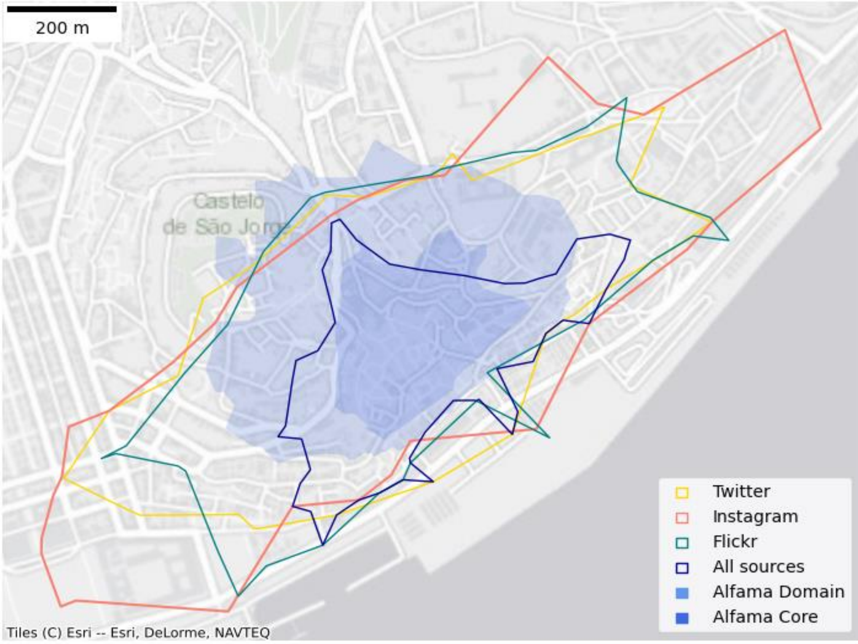

Figure 20. Comparative map of perceived and Geo-tagged boundaries for Long-term residents in Alf

Residents	Domain	Flickr	Twitter	IG	Combined
	IOU	0,580	0,534	0,384	0,394
	Overlap Km²	0.256	0.253	0.248	0.126
	Recall	0,912	0,901	0,882	0,448
	Precision	0,614	0,567	0,404	0,766
Long term	F-score	0,734	0,696	0,555	0,565
	Core	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,199	0,186	0,136	0,441
	Overlap Km²	0.0829	0.0829	0.0829	0.0756
	Recall	1,000	1,000	1,000	0,912
	Precision	0,199	0,186	0,136	0,461
	F-score	0,332	0,314	0,239	0,612

Table 7. Quantitative results between user generated content datasets Boundaries and Alfama Short-term residents ground truth.

4.3.2 Mouraria

The comparative analysis of social media platform data in delineating the perceived boundaries of the Mouraria neighborhood provides insightful metrics for both short-term and long-term residents. The results are display in the figures 21 and 22 and tables 8 and 9 .

For Short-Term Residents

In the domain region, Twitter demonstrates a significant alignment with the perceptions of short-term residents, boasting an Intersection over Union (IOU) score of 0.551 and a high recall of 0.760.Twitter’s data captures a broad area that residents associate with the Mouraria domain, although it may extend beyond the consensus boundaries. Instagram, with an exceptionally high recall of 0.940, identifies a large portion of the domain. However, its precision score of 0.447 indicates it encompasses areas that fall outside the resident-defined domain to a greater extent than Twitter. The combined sources shape, despite a low IOU value of 0.241, achieves the highest precision at 0.740, pointing to a more accurate representation within the consensus area.
Regarding the core region, Twitter's precision is notably low at 0.216, yet it nearly encapsulates the entire core area with a recall of 0.922. This implies that while Twitter's data extensively covers the core region acknowledged by residents, it is not as precise. Instagram's coverage is complete concerning the core as recognized by residents but also includes considerable areas beyond it.

For Long-Term Residents

The domain region for long-term residents again sees Twitter with a substantial IOU of 0.512 and a recall of 0.716, capturing a significant part of what long-term residents regard as Mouraria’s domain but also reaching out into areas outside of the agreed domain. Instagram presents an extensive coverage with a recall of 0.978, but its precision is only 0.475, meaning it includes more areas outside of the resident consensus compared to Twitter. The output from all data sources, while presenting a smaller overlap, aligns most accurately with the areas agreed by long-term residents.
In the core region, the metrics are generally lower. Twitter's IOU is 0.288, coupled with a recall of 0.851, showing it encompasses most of the core as identified by long-term residents. However, the precision of 0.303 suggests some imprecision in delineating specific boundaries. The combined sources output features an IOU of 0.359 and the best F-score of 0.528, indicating a

better equilibrium in capturing the core's extent while maintaining boundary accuracy.

The delineation of Mouraria by Twitter is shown to be the most representative of the neighborhood, especially when compared with the spatial agreement of long-term residents in the domain region. The selection of this delineation was based not solely on metric results, which were similar across all aspects calculated for both resident groups, but also on the historical extension of the neighborhood as described by historians. According to them, Mouraria is located between the valley of Plaça Martin Moniz and the hills of the castle and Graça neighborhood. This area was designated for the defeated Moors after the city's conquest by D. Afonso Henriques in 1147, as part of the Christian conquest over the Muslims. Therefore, the decision was made to integrate the perceived boundaries of long-term residents, which include the area between Costa de Castelo and Calçada de Santo André streets, as it better captures those described details and has moderate values in the metrics.
Another peculiarity of the results is related to the fact that neighborhood points of interest, such as the Fado Vadio Graffiti, known as a tribute to the Fado music genre and its birthplace in Mouraria, the São Cristóvão church, as ancient as it is famous, which according to the World Monuments Watch website is part of Mouraria,(Church of São Cristóvão, 2016) are not included within the perceived limits.
The exclusion of key cultural and historical landmarks from Mouraria's perceived boundaries prompts a deeper reflection on the neighborhood's vernacular ambiguity and the challenges in accurately delineating its real extent. This characteristic of Mouraria suggests a nuanced complexity in understanding its spatial identity. It underscores the necessity of considering the study area's unique characteristics even when applying Extraction methods.

The final delineation map of the historic Mouraria neighborhood, based on Twitter geotags and the perceived boundaries by long-term residents, is provided in detail in the Appendix. As a support in the results descriptions this map highlights renowned points of interest, major road arteries, and features images that capture the diverse and traditional nature defining Mouraria.

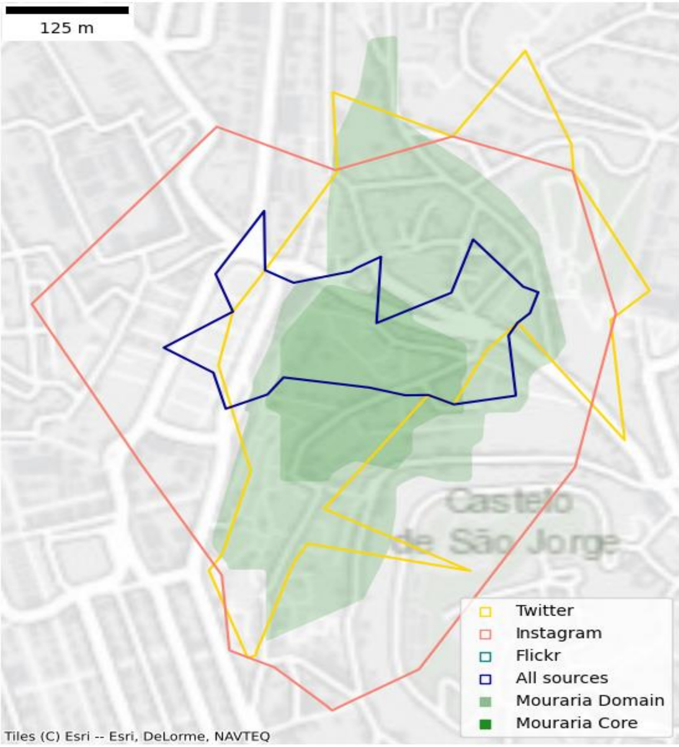

Figure 21.Comparative map of perceived and Geo-tagged boundaries for Short-term residents in Mouraria.

Residents				
	Domain	**Twitter**	**IG**	**Combined**
	IOU	0,551	0,434	0,241
	Overlap Km²	0.0950	0.1176	0.0330
	Recall	0,760	0,940	0,264
Short term	Precision	0,668	0,447	0,740
	F-score	0,711	0,605	0,389
	Core	**Twitter**	**IG**	**Combined**
	IOU	0,212	0,127	0,285
	Overlap Km²	0.0308	0.0334	0.0173
	Recall	0,922	1,000	0,519
	Precision	0,216	0,127	0,388
	F-score	0,350	0,225	0,444

Table 8. Quantitative results between user generated content datasets boundaries and Mouraria Short-term residents ground truth.

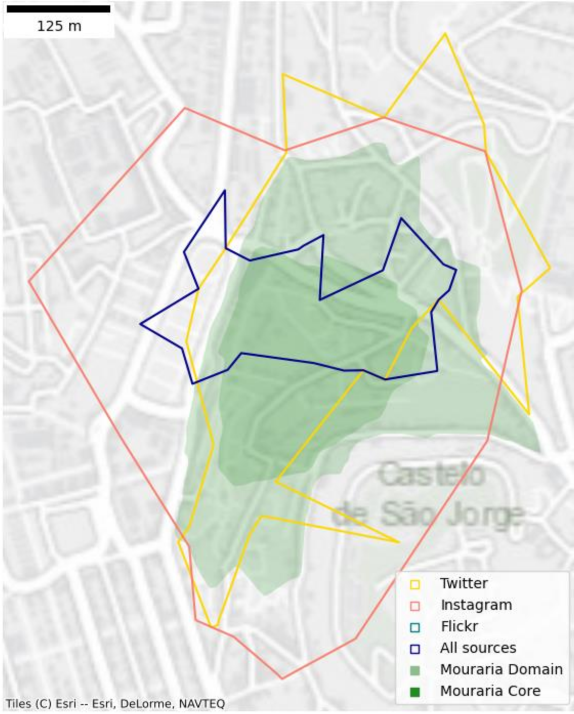

Figure 22. Comparative map of perceived and Geo-tagged boundaries for Long-term residents in Mouraria.

Residents				
	Domain	**Twitter**	**IG**	**Combined**
	IOU	0,512	0,470	0,259
	Overlap Km^2	0.0915	0.1250	0.0354
	Recall	0,716	0,978	0,277
	Precision	0,643	0,475	0,795
Long term	F-score	0,677	0,639	0,411
	Core	**Twitter**	**IG**	**Combined**
	IOU	0,288	0,193	0,359
	Overlap Km^2	0.0432	0.0507	0.0252
	Recall	0,851	1,000	0,496
	Precision	0,303	0,193	0,565
	F-score	0,447	0,323	0,528

Table 9. Quantitative results between user generated content datasets boundaries and Mouraria Long-term residents ground truth.

4.3.3 Bairro Alto

In this comparative analysis, we delve into the spatial perceptions of Bairro Alto's Domain and Core regions, as distinguished by short-term and long-term residents, using data from Flickr, Twitter, Instagram, and a combined approach. The results are display in the figures 23 and 24 and tables 10 and 11.

For Short-Term Residents

In the Domain region, Flickr's encompassing capture of the area aligns with short-term residents' recognition, even though with a precision of only 0.327, indicating extraneous area inclusion. Twitter, with a precision of 0.929 and recall of 0.487, offers a selective yet highly accurate portrayal, showing captures less but with greater fidelity. Instagram strikes a balance with recall and precision rates of 0.564 and 0.568, respectively, indicating moderate coverage and average precision. The combined data sources achieve a precision of 0.879, closely mirroring the areas short-term residents agree, despite not capturing the entirety of the Domain.
Within the Core region, Flickr's data, while all-encompassing with a recall of 1.000, suffers in precision at 0.174, suggesting it spans many areas outside the consensus. Twitter showcases a high IOU of 0.646, with recall and precision rates of 0.778 and 0.792, respectively, indicating a selective yet accurate capture of the Core. Instagram offers a broad, albeit less precise, identification, whereas the aggregate data provides a balanced representation with an IOU of 0.615, and recall and precision rates of 0.753 and 0.770, respectively.

For Long-Term Residents

Flickr again captures the entire Domain recognized by long-term residents but with low precision (0.348), indicating the inclusion of extensive non-consensual areas. Twitter's approach, with a precision of 0.910, reflects high accuracy, even covering less of the Domain area.
In the Core region, Flickr's coverage extends over the entire Core and beyond, similar to its approach in the Domain. Twitter, with its moderate recall and high precision, accurately captures the Core. The combined data sources show the highest precision for long-term residents in the Core, suggesting that an integrated approach of different data sources yields the most precise delineation.
According to (Franco, 2016), the construction of the San Roque Church in the 16th century. marked a relevant moment in the urban history of Bairro Alto, transforming it into a focal point for Lisbon's nobility. This shift led to the construction of palaces on the hill's western side, driven by a desire to be

closer to the culturally and socially esteemed priests of the Society. This historical and social background is reflected in the contemporary digital landscape, as delineations from social media platforms like Twitter, Instagram, and aggregated sources predominantly highlight the right edge over the Rua São Pedro de Alcântara, where the church is located. This area continues to attract attention, evidenced by the concentration of activity noted more than in other perceived boundaries within both resident groups.

Twitter reflects most precision in identifying the neighborhood's core, especially when contrasted with the spatial perceptions of short-term residents. This finding is intriguing when compared to other studied neighborhoods, where analysis methods struggled to accurately capture the essence of the core regions, favoring the broader domain region instead. In Bairro Alto, however, this trend is reversed, highlighting the unique urban and social fabric of the neighborhood as still centered around the historically significant San Roque Church and the main street, Rua São Pedro de Alcântara.

The final delineation map of the historic Bairro Alto neighborhood, based on Twitter geotags and the perceived boundaries by long-term residents, is provided in detail in the Appendix. As a support in the results descriptions this map highlights renowned points of interest, major road arteries, and features images that capture the vibrant nightlife and bohemian landscape.

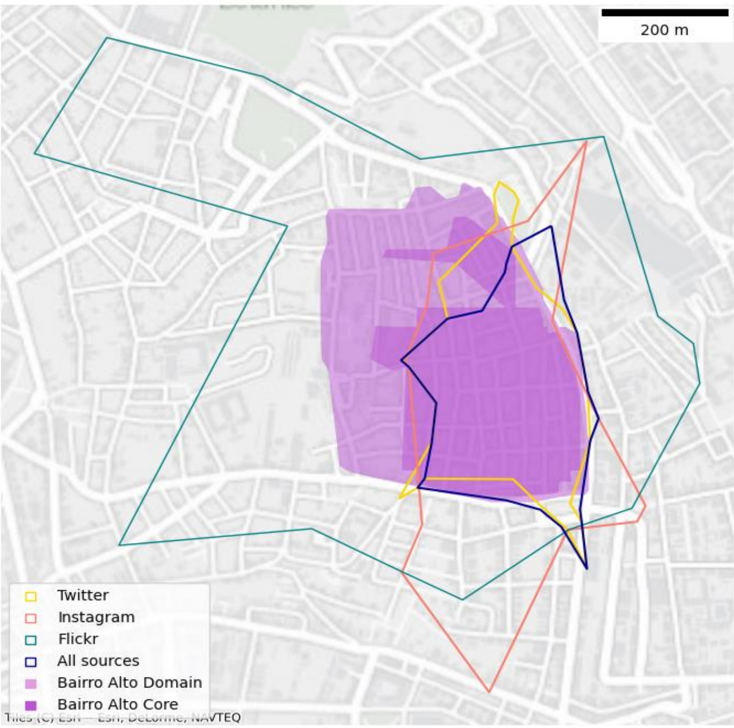

Figure 23. Comparative map of perceived and Geo-tagged boundaries for Short-term residents in Bairro Alto.

Residents					
	Domain	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,327	0,470	0,394	0,431
	Overlap Km²	0.178	0.087	0.100	0.082
	Recall	1,000	0,487	0,564	0,459
Short term	Precision	0,327	0,929	0,568	0,879
	F-score	0,493	0,639	0,566	0,603
	Core	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,174	0,646	0,452	0,615
	Overlap	0.095	0.074	0.085	0.072
	Recall	1,000	0,778	0,891	0,753
	Precision	0,174	0,792	0,479	0,770
	F-score	0,297	0,785	0,623	0,762

Table 10. Quantitative results between user generated content datasets boundaries and Bairro Alto Short-term residents ground truth.

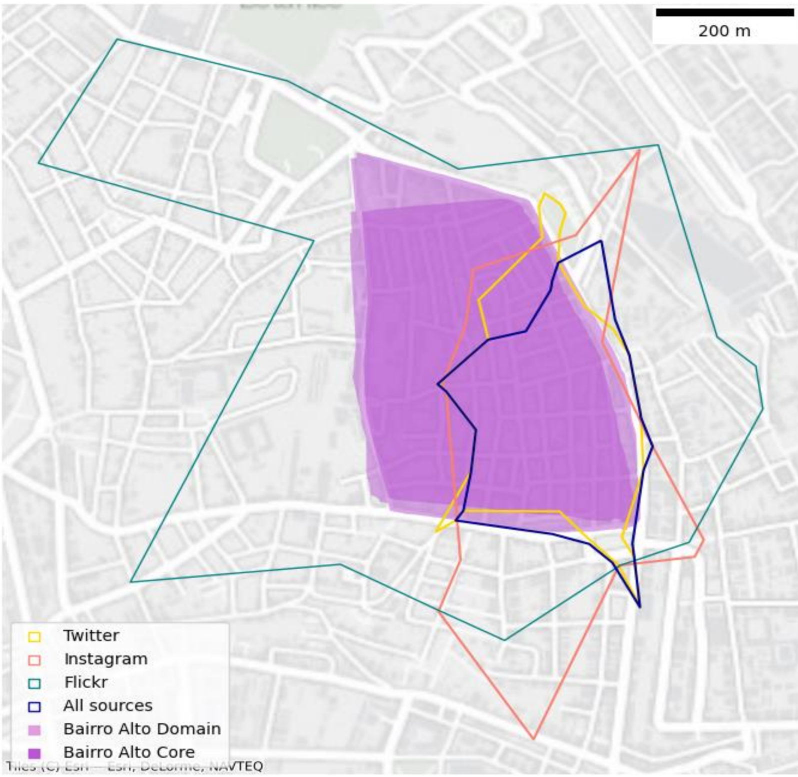

Figure 24. Comparative map of perceived and Geo-tagged boundaries for Long-term residents in Bairro Alto.

Residents					
	Domain	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,348	0,430	0,364	0,387
	Overlap	0.190	0.085	0.098	0.079
	Recall	1,000	0,449	0,516	0,416
	Precision	0,348	0,910	0,552	0,848
Long term	F-score	0,516	0,601	0,533	0,558
	Core	**Flickr**	**Twitter**	**IG**	**Combined**
	IOU	0,281	0,438	0,343	0,362
	Overlap	0.153	0.075	0.084	0.065
	Recall	1,000	0,491	0,551	0,427
	Precision	0,281	0,804	0,476	0,704
	F-score	0,438	0,610	0,511	0,532

Table 11. Quantitative results between user generated content datasets and Bairro Alto Ground truth.

.

Chapter 5

5. Final Discussion

5.1 Answering Research Questions

Based on previous results and supported by literature from research with similar approaches, these are the answers to the research question of this study.

How do residents spatially define different historic neighborhoods in the city of Lisbon?

We can assert that Lisbon's city residents spatially define historic neighborhoods based on their length of residency. The division between short-term and long-term residents revealed a spatial extension and social details associated or not associated with the obtained surfaces varied according to the residency duration due to the creation of ties with each neighborhood, as studied by (Tang et al., 2021)and (Hernández et al., 2007). Residents define the spatial embedding of everyday activity spaces related to the local availability of outlined points of interest, as it is explained in (Westerholt et al., 2022) study. Other factors influencing the delineation of perceived boundaries are the street network and the structural age of a neighborhood, due to orientation patterns that affect the cognitive representation of urban form. This is particularly evident during the drawing of mental maps linling physical and social factors, as seen in the results of the perceived boundaries of the three historic neighborhoods and mentioned in (Westerholt et al., 2022) and (Tang & Painho, 2023) studies. Finally, landform features are crucial in inference-based neighborhood mapping methods, as is described in (Tang & Painho, 2023) study, with notable factors including Lisbon's topography of slopes and hills.

What are the similarities and differences regarding neighborhood boundaries between primary and secondary data as well as between the different sources of secondary data?

Mental maps derived directly from residents via web-based surveys constitute the primary data, supplemented by points of geo-tagged online activity from diverse sources of user-generated content (UGC) as secondary data. These surfaces were further analyzed in conjunction with literature on studies that also integrate both data types alongside historical and

social contexts. In terms of similarities between data sources, the delineated boundaries reveal that these areas not only mirror user practices and perceptions but also underscore the diversity of preferences, mobility, interactions and neighborhood identity and behavior within every study area, (Bae & Montello, 2018). Nonetheless, the delineation of check-in clusters from UGC sources such as Twitter, Instagram, or Flickr showed discrepancies. While geotag surfaces in Alfama and Mouraria sometimes captured the ground truth, their accuracy was not consistent. Conversely, in Bairro Alto, geotags precisely captured a segment of the ground truth, highlighting the variable nature of public life development across areas. This resulted in distinct clusters in more vibrant areas, whereas less central areas displayed reduced social activity, as noted by (Chugunov et al., 2019).
Huang's findings, utilizing the same kind of data sources as this study, advocate for the combination of data sources as a methodological strength that deepens our understanding of specific areas, (Huang et al., 2021). Moreover, when sufficient resources are available, it is possible to create composite delineations that closely represent the authentic character of a region, aligning with Lynch's suggestions, (Lynch, 1964).

What are the spatial footprints of historic neighborhoods according to different sources of geo-tagged user-generated content?

According to (Hollenstein & Purves, 2010), virtual interactions shed light on the perception of public spaces, with Twitter particularly emphasizing the relevance of everyday life venues for communities. Our research indicates that in Alfama, Twitter's spatial delineation aligns with patterns observed across other social networks, exhibiting moderate metric values, yet it is not deemed the most representative. In contrast, Twitter's delineations in Mouraria more closely match the areas recognized by long-term residents, albeit not with high accuracy. Notably, in Bairro Alto, Twitter effectively delineates the core region, reflecting a high level of social integration among students, residents, and tourists drawn to local amenities.
Hollenstein & Purves analysis further reveals that Instagram is associated with architectural landmarks and tourist attractions. However, despite Alfama's appeal as an 'Instagrammable' neighborhood, the delineations from Instagram were extensive yet imprecise. Delineation efforts from Instagram to represent Mouraria faced challenges due to the dispersion of the dataset, whereas in Bairro Alto, Instagram delineated areas of high social activity, mirroring Twitter's findings. Flickr, catering to photography enthusiasts, delineated Alfama with high precision, suggesting that users have a clear understanding of specific places, their locations, and extents.

The combination of data from all sources successfully identified the core regions of the three neighborhoods, an achievement not mirrored by the individual social networks. These combined data surfaces were compact and localized to the neighborhoods' cores, consistent with patterns observed in comparisons with data from both short- and long-term residents. This underscores how user activities and intentions reflect the diverse spatial delineation of historic neighborhoods, supported by clustering methods that unveil boundaries closely approximating reality based on interactions as (Tang et al., 2022).

How reliable is the use of geo-tagged online activity in inferring residents' opinions on neighborhood boundaries?

Based on the outcomes of this research, which obtained delineations derived from geo-tagged online activity, it is suggested that while the delineations are not entirely precise, these data source can indeed infer residents' opinions on neighborhood boundaries. The use of this data not only captures spatial patterns but also taps into the semantics of cognitive regions, as demonstrated by (Gao et al., 2017b) research. Li's study supports UGC reliability by showing that openly available data, such as collected tweets in her study, can effectively outline administrative boundaries and major roads in the United States, particularly in areas of high population density, (Li et al., 2013). Furthermore, (Huang et al., 2021) argues that geolocated social media data is both timely and relevant, providing a contemporary supplement to more traditional methods of characterizing urban public spaces. Overall, the integration of social media analytics into urban studies offers a resonating with the perceptions of residents.

5.2 Limitations

One of the most challenging aspects of conducting this research was related to one of the project's objectives: to retrieve representative boundaries from various sources of geo-tagged online activity using neighborhood names as keywords. Various filters were implemented to extract specific data successfully. However, the extraction of additional aspects could help expand the datasets and generate delineations that might yield improved results. This specificity in data retrieval resulted in the absence of a dataset from the social network Flickr for the Mouraria neighborhood. Consequently, it is unclear if the variability in the distribution patterns observed in the datasets from other social networks would be mirrored in Flickr.

5.3 Future Scope

Future research could benefit from expanding the extraction of textual attributes related to the specific economic, environmental, and social characteristics of each neighborhood. Efforts could include exploring diverse methods for unifying perceived boundaries, such as boundary aggregation or radial averaging due to, as mentioned in the method comparison study by (Dalton & Hurrell, 2023). Additionally, experimenting with various clustering algorithms like spectral clustering or Kernel Density Estimation could prove valuable. Incorporating insights from ethnographic studies, like those conducted by (Baptista et al, 2018), which reveal the less positive social realities within the three neighborhoods, would be particularly insightful. These realities, including issues such as marginalization and social exclusion, are reflected in the discourse on social media and could add depth to this study's findings. If this approach proves successful, it could be applied to other historic neighborhoods in the city, each distinguished by its unique characteristics, thereby enhancing our understanding of urban social dynamics.

Chapter 6

6. Conclusions

This study embarked on the intricate challenge of mapping out the vernacular boundaries of Lisbon's historic neighborhoods: Alfama, Mouraria, and Bairro Alto. By blending residents' insights with geo-tagged data, alongside methods and tools for extracting perceptions, we've unveiled the true potential of merging cognitive mapping with user-generated content. This method has proven effective in delineating the spatial confines of these storied areas, illuminating the rich tapestry of history and community life that's interwoven with the urban fabric of Lisbon.

The findings shed light on the nuanced ways digital platforms and social interactions contribute to our grasp of urban boundaries. Alfama's representation through Flickr's lens emerged as notably precise, underscoring the impact of visual platforms in capturing the essence of locales, thus emphasizing the significant role of visual perceptions in boundary delineation. In Mouraria, Twitter's alignment with long-term residents' understanding highlighted the platform's capability to capture the neighborhood's essence via social media dynamics. Meanwhile, Bairro Alto's detailed mapping via Twitter revealed the influence of historical and social underpinnings in shaping urban borders.

Addressing the research questions, it's insightful to note how residents spatially define different historic neighborhoods through various lenses, be it the length of residence, the orientation of street networks, or the spatial extent of areas. A composite approach, which amalgamates multiple data sources, affords a richer, more accurate representation of the core regions of vernacular boundaries. Additionally, the findings advocate that social media analytics can serve as a reliable index for perceived city images, with platforms like Instagram elucidating public perceptions tied to tourist attractions and landmarks.

The implications of this research reverberate through the realms of urban planning and policymaking, championing an approach that more closely resonates with the lived experiences and perceptions of city inhabitants. A deeper understanding of subjective neighborhood delineations can significantly refine strategies in tourism management and resource allocation, especially within highly marginalized areas. This thesis contributes to the enrichment of vernacular geography by offering new insights into the modern urban fabric, shaped by a combination of social interactions and digital footprints.

The successful integration of spatial and social dimensions showcased in this study underscores the efficacy of employing diverse methodologies for data extraction, analysis, clustering, and delineation. This multifaceted approach not only enhances our comprehension of urban spaces but also lays down a comprehensive framework for navigating the intricate interplay between physical geography and social constructs in defining neighborhood boundaries.

7. Bibliographic References

Anacta, V. J. A., Humayun, M. I., Schwering, A., & Krukar, J. (2017). Investigating Representations of Places with Unclear Spatial Extent in Sketch Maps. In A. Bregt, T. Sarjakoski, R. Van Lammeren, & F. Rip (Eds.), *Societal Geo-innovation* (pp. 3–17). Springer International Publishing. https://doi.org/10.1007/978-3-319-56759-4_1

Arribas-Bel, D., Garcia-López, M.-À., & Viladecans-Marsal, E. (2021). Building(s and) cities: Delineating urban areas with a machine learning algorithm. *Journal of Urban Economics*, *125*, 103217. https://doi.org/10.1016/j.jue.2019.103217

Bae, C., & Montello, D. (2018). Representations of an Urban Ethnic Neighbourhood: Residents' Cognitive Boundaries of Koreatown, Los Angeles. *Built Environment*, *44*, 218–240. https://doi.org/10.2148/benv.44.2.218

Baptista, L. V., Nofre, J., & Jorge, M. do R. (2018). Mobilidade, Cidade e Turismo: Pistas para analisar as transformações em curso no centro histórico de Lisboa. *Sociologia: Revista da Faculdade de Letras da Universidade do Porto*, 14–32.

Brown, G., Raymond, C. M., & Corcoran, J. (2015). Mapping and measuring place attachment. *Applied Geography*, *57*, 42–53. https://doi.org/10.1016/j.apgeog.2014.12.011

Chugunov, A., Misnikov, Y., Roshchin, E., & Trutnev, D. (Eds.). (2019). *Electronic Governance and Open Society: Challenges in Eurasia: 5th International Conference, EGOSE 2018, St. Petersburg, Russia, November 14-16, 2018, Revised Selected Papers* (Vol. 947). Springer International Publishing. https://doi.org/10.1007/978-3-030-13283-5

Church of São Cristóvão. (n.d.). World Monuments Fund. Retrieved February 23, 2024, from https://www.wmf.org/project/church-s%C3%A3o-crist%C3%B3v%C3%A3o

Cocola-Gant, A., & Gago, A. (2021). Airbnb, buy-to-let investment and tourism-driven displacement: A case study in Lisbon. *Environment and Planning A: Economy and Space*, *53*(7), 1671–1688. https://doi.org/10.1177/0308518X19869012

Coulton, C. J., Jennings, M. Z., & Chan, T. (2013). How Big is My Neighborhood? Individual and Contextual Effects on Perceptions of Neighborhood Scale. *American Journal of Community Psychology*, *51*(1–2), 140–150. Scopus. https://doi.org/10.1007/s10464-012-9550-6

Da Costa, A. F. (2008). *Sociedade de Bairro Dinâmicas Sociais da Identidade Cultural* (2th ed.). Celta editora.

Dalton, N. S., & Hurrell, M. (2023). Methods for neighbourhood Mapping, boundary agreement. *Environment and Planning B: Urban Analytics and City Science*, *50*(2), 401–415. https://doi.org/10.1177/23998083221115195

Deng, Y. (2016). Challenges and complications in neighborhood mapping: From neighborhood concept to operationalization. *Journal of Geographical Systems*, *18*(3), 229–248. https://doi.org/10.1007/s10109-016-0232-z

Edelsbrunner, H., Kirkpatrick, D., & Seidel, R. (1983). On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, *29*(4), 551–559. https://doi.org/10.1109/TIT.1983.1056714

Estevens, A., Cocola-Gant, A., Calvo, D. M., & Matos, F. (2019). As artes e a cultura nas práticas hegemónicas e alternativas na cidade de Lisboa. O caso do Largo do Intendente. *Forum Sociológico. Série II*, *35*, Article 35. https://doi.org/10.4000/sociologico.7826

Evans, A. J., & Waters, T. (2007). Mapping vernacular geography: Web-based GIS tools for capturing "fuzzy" or "vague" entities. *International Journal of Technology, Policy and Management*, *7*(2), 134. https://doi.org/10.1504/IJTPM.2007.014547

Franco, A. (2016). *Caminhar por Lisboa 7 percursos pelas histórias e segredos da cidade* (first). Porto Editora.

Gao, S., Janowicz, K., Montello, D. R., Hu, Y., Yang, J.-A., McKenzie, G., Ju, Y., Gong, L., Adams, B., & Yan, B. (2017a). A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science*, *31*(6), 1245–1271. https://doi.org/10.1080/13658816.2016.1273357

Gao, S., Janowicz, K., Montello, D. R., Hu, Y., Yang, J.-A., McKenzie, G., Ju, Y., Gong, L., Adams, B., & Yan, B. (2017b). A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science*, 1–27. https://doi.org/10.1080/13658816.2016.1273357

Glebova, M. (n.d.). *Town and Gown: Visualising University Neighbourhoods as Places within the Urban Environment*.

Hernández, B., Carmen Hidalgo, M., Salazar-Laplace, M. E., & Hess, S. (2007). Place attachment and place identity in natives and non-natives. *Journal of Environmental Psychology*, *27*(4), 310–319. Scopus. https://doi.org/10.1016/j.jenvp.2007.06.003

Hollenstein, L., & Purves, R. S. (2010). Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science*, *1*(2010), 21–48. Scopus. https://doi.org/10.5311/JOSIS.2010.1.3

Huang, J., Obracht-Prondzynska, H., Kamrowska-Zaluska, D., Sun, Y., & Li, L. (2021). The image of the city on social media: A comparative study using "Big Data" and "Small Data" methods in the Tri-City Region in Poland. *Landscape and Urban Planning*, *206*, 103977. https://doi.org/10.1016/j.landurbplan.2020.103977

Langens, J. (2022). Mouraria crossroads of cultures as old as the nation. In *10 Lisbon Stories*. casadasletras.

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography* *and Geographic Information Science*, *40*(2), 61–77. https://doi.org/10.1080/15230406.2013.777139

Lynch, K. (1964). *The Image of the City*. MIT Press.

Meinig, D. W. (1965). The Mormon Culture Region: Strategies and Patterns in the Geography of the American West, 1847–1964. *Annals of the Association of American Geographers*, *55*(2), 191–219. https://doi.org/10.1111/j.1467-8306.1965.tb00515.x

Nofre, J. (2020). *Nofre, J. (2020). Party tourism, pub crawls and crazy nights in Bairro Alto, Lisbon: Celebrating cosmopolitanism in a racialized heterotopic pleasurescape. IS Working Papers, Serie III, No. 87: 1-13.*

Oliveira, N. (2019). A Lisboa cosmopolita e o fascínio da diversidade. *CIDADES, Comunidades e Territórios*, 115–128. https://doi.org/10.15847/citiescommunitiesterritories.dec2019.039.art04

Salvatore, C., Biffignandi, S., & Bianchi, A. (2021). Social Media and Twitter Data Quality for New Social Indicators. *Social Indicators Research*, *156*(2), 601–630. https://doi.org/10.1007/s11205-020-02296-w

Tang, V., Acedo, A., & Painho, M. (2021). Sense of place and the city: The case of non-native residents in Lisbon. *Journal of Spatial Information Science*, *23*, Article 23. https://doi.org/10.5311/JOSIS.2021.23.165

Tang, V., & Painho, M. (2023). Exploring the relationships between perceived neighborhood boundaries and street network orientation. *Transactions in GIS*, *27*(3), 877–899. https://doi.org/10.1111/tgis.13058

Tang, V., Puri, J., & Painho, M. (2022). Mapping a historic neighbourhood through user-generated content: The case of Alfama, Lisbon (Portugal). *AGILE: GIScience Series*, *3*, 1–8. https://doi.org/10.5194/agile-giss-3-63-2022

Tourism industry eclipses pre-pandemic levels in 2023—Products Eurostat News—Eurostat. (n.d.). Retrieved February 5, 2024, from https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20240116-1

Vernon, H. (2023, February 27). *2022, el mejor año turístico de la historia de Portugal*. Benoit Properties. https://www.benoitproperties.com/es/news/2022-marks-best-ever-year-for-portugal-tourism/

Westerholt, R., Acedo, A., & Naranjo-Zolotov, M. (2022). Exploring sense of place in relation to urban facilities – evidence from Lisbon. *Cities*, *127*, 103750. https://doi.org/10.1016/j.cities.2022.103750

Zelinsky, W. (1980). North America's Vernacular Regions. *Annals of the Association of American Geographers*, *70*(1), 1–16.

Appendix

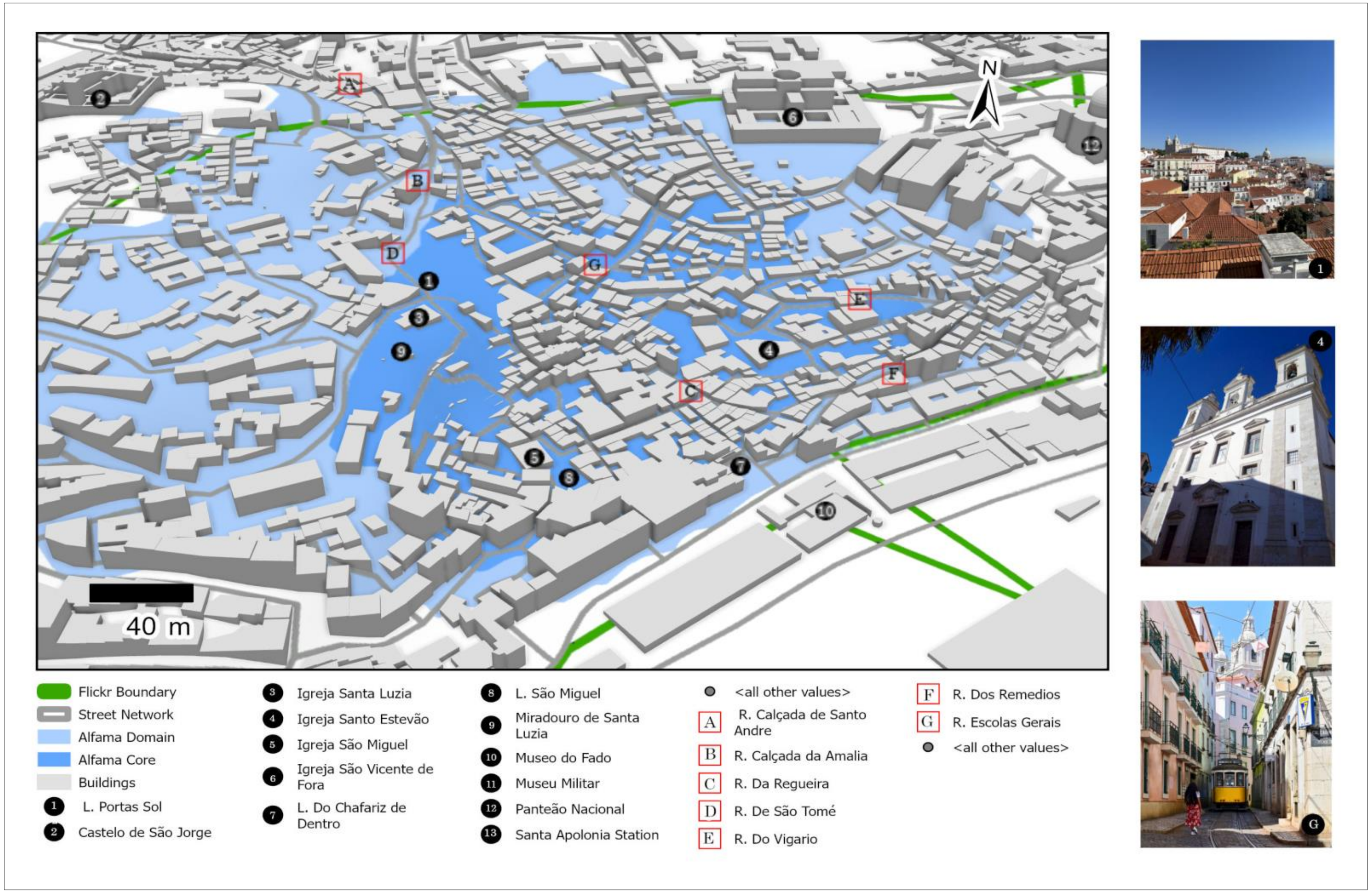

Figure 25. Map of the Historic Alfama Neighborhood Delineation Based on Flickr Geotags and Perceived Boundaries by Long-Term Residents

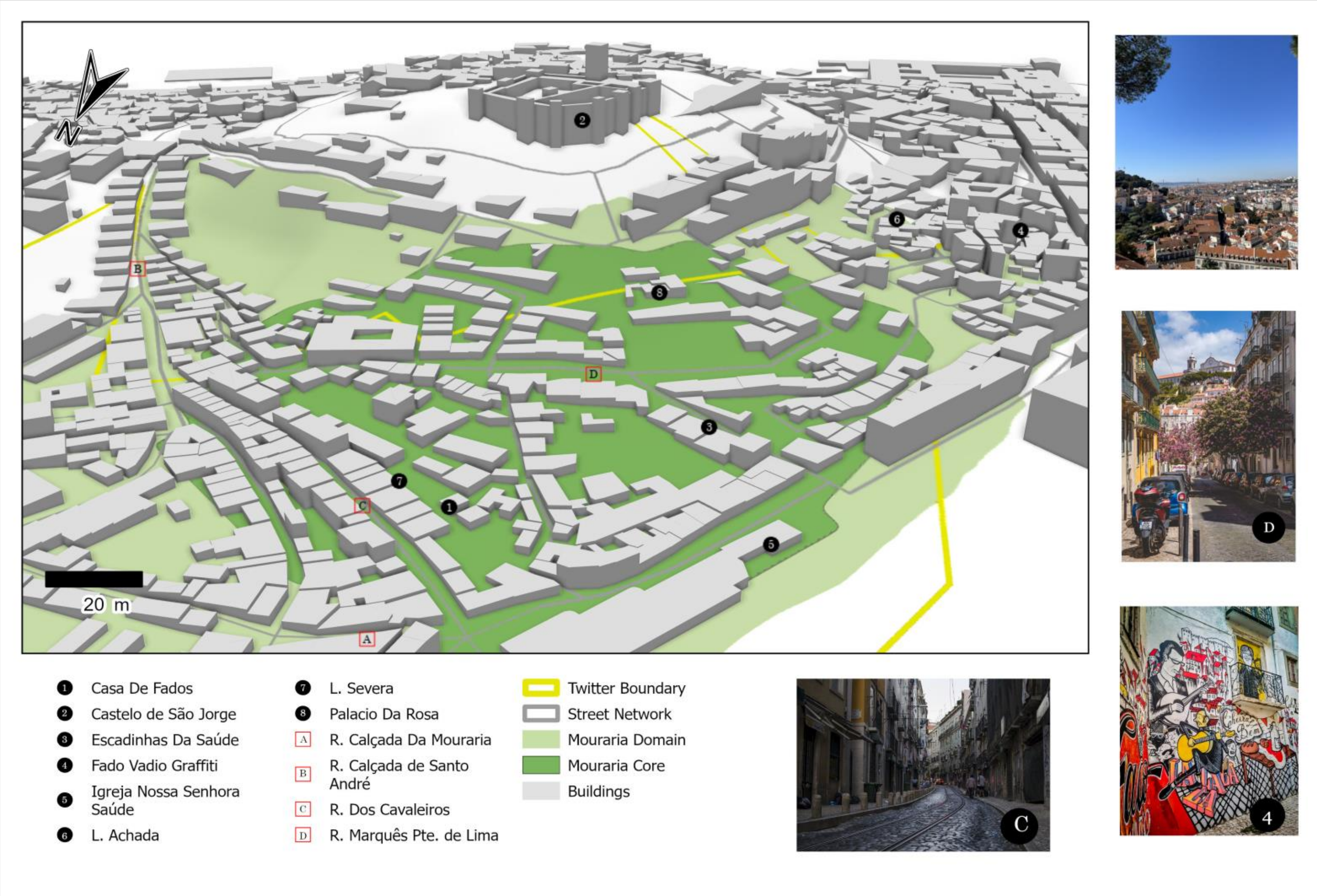

Figure 26. Map of the Historic Mouraria Neighborhood Delineation Based on Twitter Geotags and Perceived Boundaries by Long-Term Residents

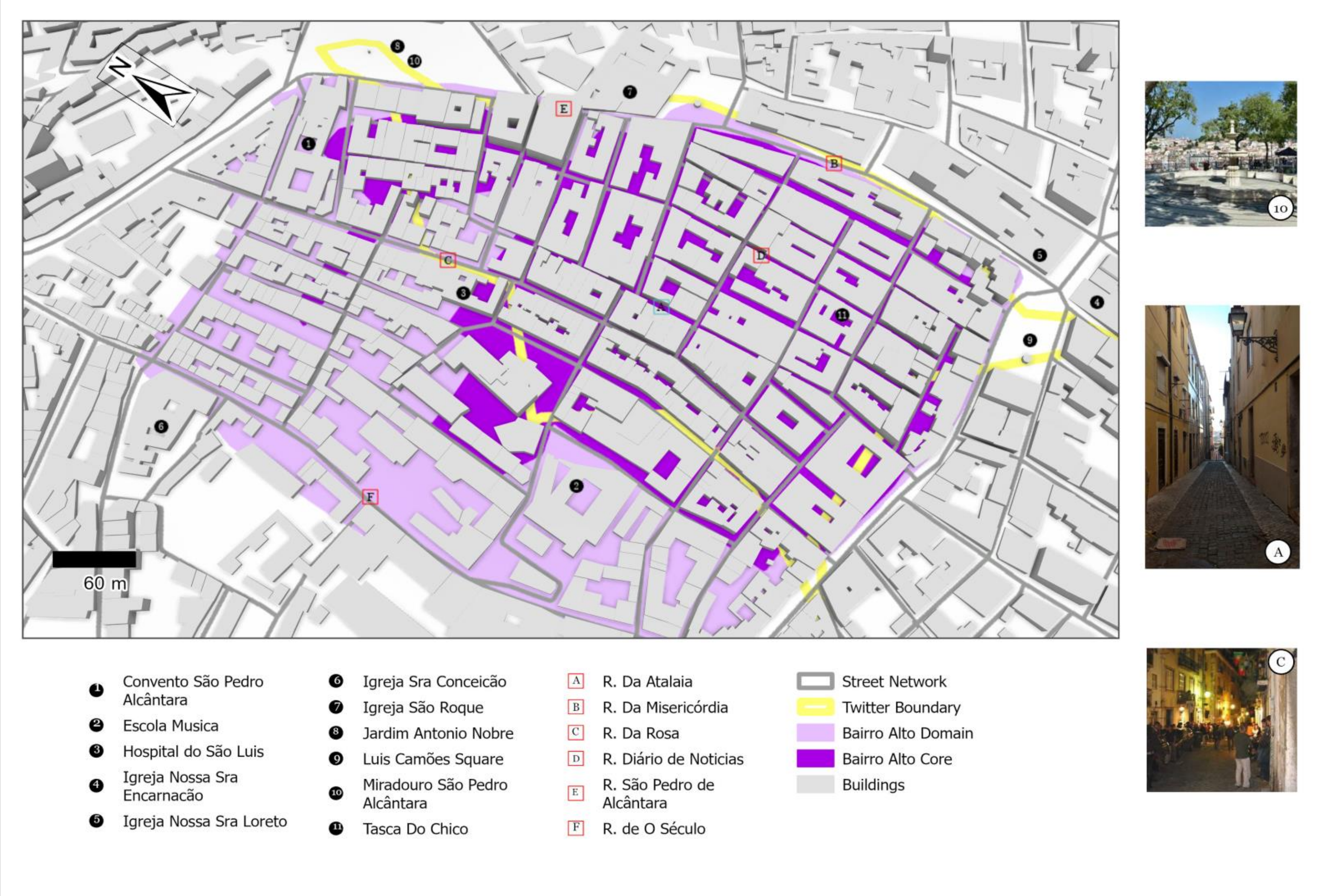

Figure 27. Map of the Historic Bairro Alto Neighborhood Delineation Based on Twitter Geotags and Perceived Boundaries by Long-Term.Residents

Masters
Program
in **Geospatial**
Technologies

Supported by:
Education and Culture
ERASMUS MUNDUS