

## Mapping Chinese Diasporic Heritage from Social Media Data with Spatial and Semantic Networks

Zhou, Y.; Bai, N.; Spoormans, L.G.K.; Pereira Roders, A.

**Publication date**

2025

**Document Version**

Final published version

**Citation (APA)**

Zhou, Y., Bai, N., Spoormans, L. G. K., & Pereira Roders, A. (2025). *Mapping Chinese Diasporic Heritage from Social Media Data with Spatial and Semantic Networks*. Paper presented at 19th International Conference on Computational Urban Planning and Urban Management (CUPUM 2025), London, United Kingdom.

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Mapping Chinese Diasporic Heritage from Social Media Data with Spatial and Semantic Networks

Yan Zhou<sup>\*1</sup>, Nan Bai<sup>†1</sup>, Lidwine Spoormans<sup>‡1</sup> and Ana Pereira Roders<sup>§1</sup>

<sup>1</sup>Faculty of Architecture and the Built Environment, Delft University of Technology

February 10, 2025

## Summary

Diasporic communities and their heritage are vital in shaping cultural diversity in urban planning and management. However, diasporic heritage management overlooked cross-geographical networks that constitute cultural significance. This study develops a workflow to explore diasporic heritage’s spatial and semantic networks. Global social media data about Chinatowns on Flickr is collected and analysed through Named Entity Recognition, and Spatial and Semantic Network Analysis. Findings reveal the diasporic flows based on places frequently recognised as origin and destination of Chinese diasporic heritage, and the distribution of places, zooming in from worldwide to the Netherlands. Semantic networks in different Dutch cities are compared.

**KEYWORDS:** Diasporic heritage, Chinatown, Social media data, Spatial network, Semantic network

## 1 Introduction

Cross-cultural migration and diaspora have been a historical phenomenon, and are experiencing more complex and rapid mobilities in contemporary times (J. Lucassen and L. Lucassen 2009; Castles 2007). Diasporic heritage is created and maintained by diasporic communities. However, research on identifying, assessing, and managing diasporic heritage remains limited (Dellios and Henrich 2020). Although analysis of origin-destination interrelations is indispensable in migration studies (Shuval 2000), most studies in heritage management overlooked these interrelations (Byrne 2016). Understanding the cross-geographical networks of heritage attributes and diasporic communities is essential for a comprehensive assessment of cultural significance, which also helps to strengthen the cultural diversity in urban planning and management (Fincher et al. 2014).

Heritage can be understood by how it is shared on social media. The distribution of post geotags can reflect the places of heritage interests (Ginzarly and Srouf 2021; Bai et al. 2023). The entities (e.g.

---

<sup>\*</sup>yan.zhou@tudelft.nl

<sup>†</sup>n.bai@tudelft.nl

<sup>‡</sup>l.g.k.spoormans@tudelft.nl

<sup>§</sup>a.r.Pereira-Roders@tudelft.nl

nouns, placenames) mentioned in the posts can reference heritage attributes and relevant locations (Lai 2019; Ginzarly, Pereira Roders, and Teller 2019). For migration-related studies, social media offers timely and granular data, allowing for immediate and precise analysis (Danchev and Porter 2021).

This study aims to map the diasporic flows and visualise the networks of noun entities based on social media data. By understanding the cross-geographical characteristics, this study can contribute to cultural significance assessment of diasporic heritage.

## **2 Methodology**

The process of data collecting, processing, and network analysis is illustrated in **Figure 1**.

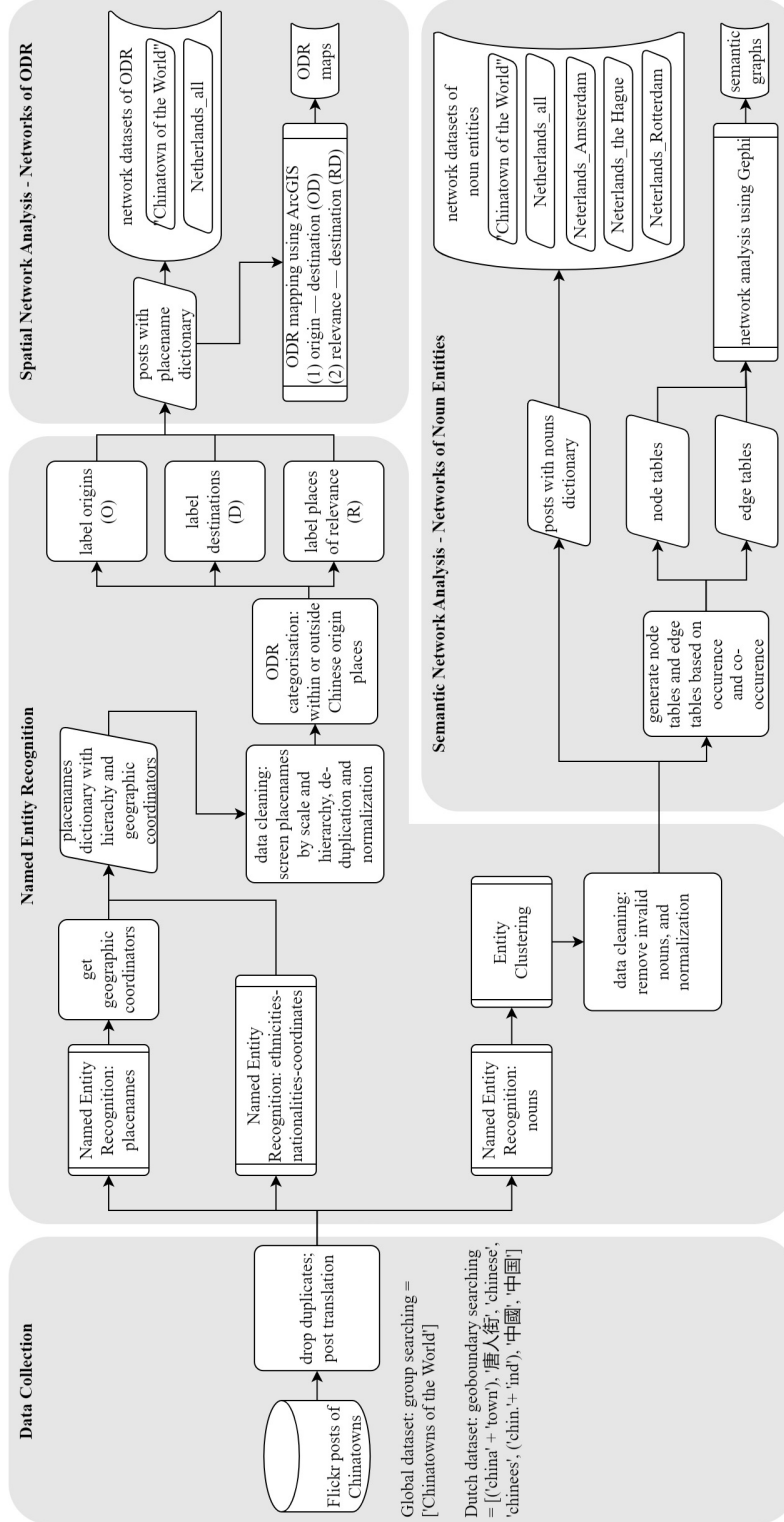


Figure 1: Workflow of this study with steps of data collection, named entity recognition, and spatial and semantic analysis

## 2.1 Data Collection

A photo-sharing social media platform, Flickr, is used as the data source. Two post datasets are collected: the global Chinatowns and the Chinese diasporic heritage in the Netherlands.

18,975 posts are exhausted from the Flickr group “Chinatown of the World”<sup>1</sup>, and 3,632 posts within the Netherlands’ geographical boundary are collected with keywords. i.e., “(‘china’ + ‘town’), ‘唐人街’, ‘Chinese’, ‘Chinese’, (‘chin.’+ ‘ind’), ‘中國’, ‘中国’ ”.

Photo IDs, geotags (lat, lon, if applicable), and text information (title, description, and hashtags) are retrieved and translated into English<sup>2</sup>. Unrelated posts are removed based on user IDs and keywords.

## 2.2 Named Entity Recognition to Extract ODR and Noun Entities

BERT NER package<sup>3</sup> is used to extract placenames and noun entities. Recognised placenames are cleaned using GeoNames API. Only placenames with administrative hierarchies are kept<sup>4</sup> for coordinates retrieving. Noise is reduced through deduplication and normalisation.

To map diasporic flows, an ODR network is built on the OD (origin and destination) network originated from Migration and Transportation studies (Yuhang Gu and Mocnik 2023). Another category, “places of relevance” (R), is introduced to identify placenames not directly origins or destinations. The posts are then labelled with the following rule:

1. Placenames inside Chinese original places (i.e., Mainland China, Hong Kong, Macau, and Taiwan) are labelled “origin” (O).
2. The geotags of the posts are labelled “destination” (D).
3. Placenames outside Chinese original places are labelled “places of relevance” (R).

In addition, exploring the nouns of ethnicities in diasporic heritage posts can expand the ODR network, and discover the potential stakeholders of local heritage sites. Ethnicities are screened for each post according to a dataset of ethnicity - nationality (country name) - coordinates information<sup>5</sup> to roughly trace back other related communities and countries of “origin” (O), and “relevance” (R).

---

<sup>1</sup>This group specifically focuses on documentation and pan-global reflection of all Chinatowns of the World. Strict selection rules from the group ensure that posts always meet the theme of “Chinatown”. See group description: <https://www.flickr.com/groups/2307218@N24/>

<sup>2</sup>Google Cloud Translate API is used for translation

<sup>3</sup>A fine-tuned BERT model for Named Entity Recognition: “dslim/bert-base-NER”. See <https://huggingface.co/dslim/bert-base-NER> (accessed on 30/12/2024)

<sup>4</sup>PCLI (country), ADM1 (state/province), PPLA (city), PPLC (country capital), PPLX (district/neighbourhood)

<sup>5</sup>This dataset merged two open-source datasets: a dataset of worldwide nationality (see <https://github.com/nicolanrizzo/nationalitylist>, accessed on 29/12/2024) and a dataset of country coordinates (see <https://github.com/cristiroma/countries>, accessed on 29/12/2024). The merged dataset provides the following structure for matching the country name and the geographic coordinates to the ethnicity, e.g. Chinese: China: (36.7145744, 103.558192), Americans: USA: (37.66895362, -102.3925645), Singaporean: Singapore: (1.33873261, 103.8332356)

Noun entity groups are extracted from the posts. The noise is reduced by semantic clustering with HDBSCAN algorithm.

### 2.3 Spatial and Semantic Network Analysis

Chinese diasporic heritage’s global and Dutch ODR networks are visualised in ArcGIS. Semantic networks of noun entities are plotted using Gephi.

In the semantic networks, nodes are weighted by the entity cluster sizes. Edges represent co-occurrences of cluster groups within individual posts, weighted by the frequency of co-occurrence. Comparisons among subnetworks of Amsterdam, the Hague, and Rotterdam are made to examine localised distinctions.

## 3 Results

### 3.1 Spatial Network of Diasporic Flows

Three areas have the largest number of Chinatown posts: North America (mainly in Canada and the USA), Europe (mainly in the UK, Netherlands, France, and Italy), and Southeast Asia (mainly in Malaysia and Singapore). Chinatowns in Japan and Australia are also hotspots of destinations. Compared to Chinatowns that are already well-known, like in San Francisco (USA), and Vancouver (Canada), some Chinatowns with fewer previous studies arouse attention, such as those in Honolulu (USA), Havana (Cuba), Buenos Aires (Argentina), Ho Chi Minh City (Vietnam), Ipoh (Malaysia), and Honiara (Solomon Islands). Hotspots of origin are seen in Guangdong and Taiwan. Many OD links connect Guangdong to Australia, and Taiwan to Japan. A few posts refer to Thailand and Myanmar as places of relevance, while no destination posts are found in these countries (**Figure 2**).

In the global dataset, only 7 posts in the Hague and 10 posts in Amsterdam are mentioned from the Netherlands. However, in the Dutch dataset, more posts are found in destinations like Amsterdam (727), Rotterdam (465), and the Hague (345). Other hotspots scatter in other cities, such as Tilburg, Arnhem, Utrecht, Eindhoven, and Groningen. More origins in the Dutch dataset locate in Shanghai and Hong Kong, and relevant countries like Indonesia, Surinam, Japan and India (**Figure 2**). Such differences show that the visibilities of Chinese diasporic heritage might vary from global to national scales.

As shown in the examples (**Figure 3**), Chinese diasporic heritage ODR networks reveal interrelations of cultural significance between the ODR places, such as migration histories, community interactions, aesthetic and term relations, and changes in time.

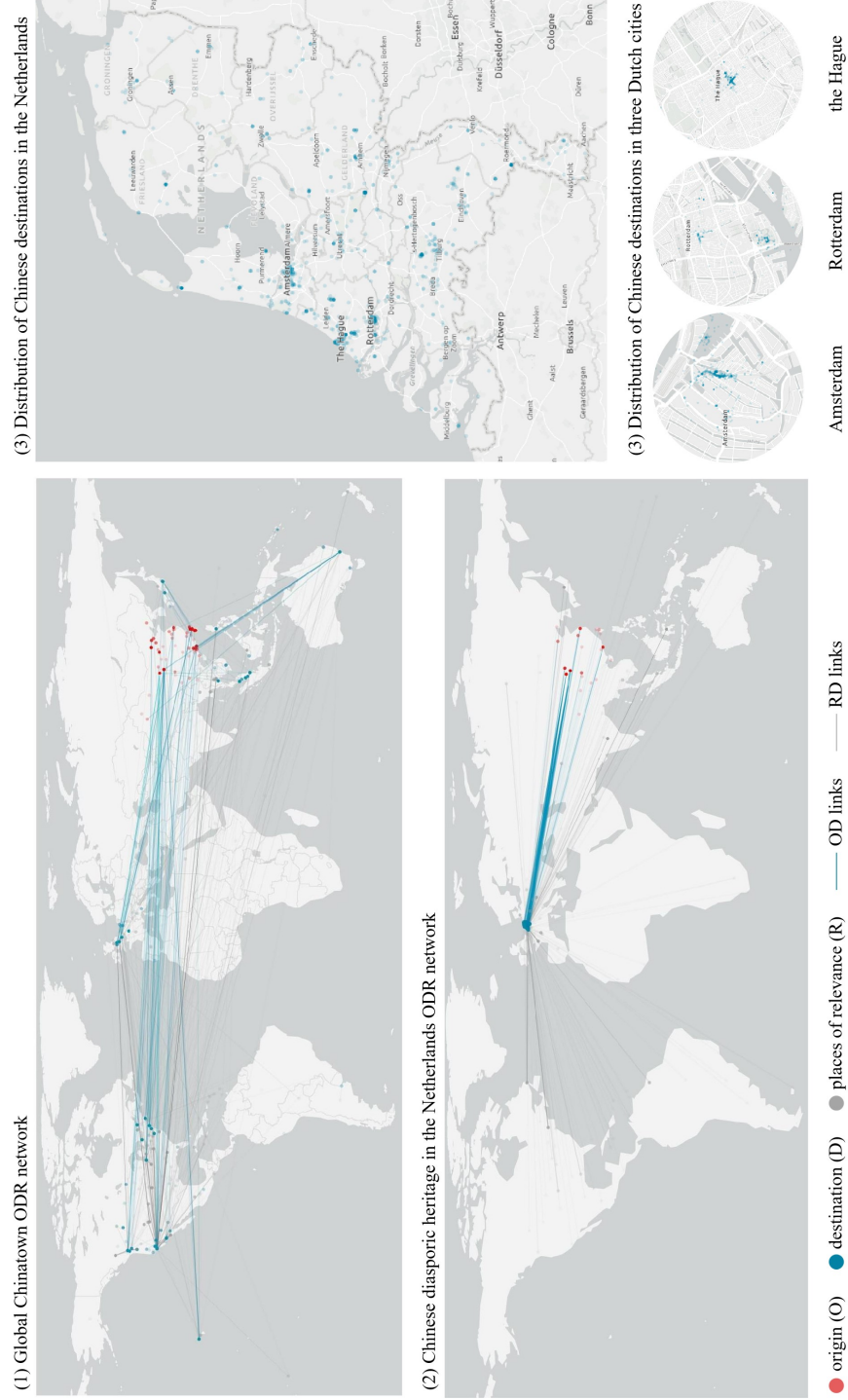


Figure 2: Mapping of diasporic flows and distributions respectively in the global and Dutch datasets. The points of O, D, and R are plotted in red, blue, and grey. The blue lines represent the OD links, and the grey lines represent the RD links.

| Examples of Global dataset                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                  |                                                                                                   |                                    |                                               |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|------------------------------------|-----------------------------------------------|
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | O                                                                                                | D                                                                                                 | R                                  | ODR Interrelations                            |
| (1) "When <b>Chinese</b> settlers in <b>Sydney</b> decided to build a temple dedicated to the <b>Chinese</b> folk hero and god, Kwung Ti, they prayed for a sign showing where the temple should be built. This led them to a market garden in Francis Street (now Edward Street) in <b>Glebe</b> which they purchased from the owner for £325. The temple was funded by immigrants from the area known as <b>Sze Yup</b> in the province of Kwongtung, <b>China</b> . After 100 years, the <b>Sze Yup</b> Temple remains a sacred place to the <b>Chinese</b> community. Worshippers now include new arrivals from <b>Hong Kong</b> , <b>Malaysia</b> , <b>Singapore</b> , <b>Taiwan</b> , <b>China</b> and, in particular, <b>Indo-China</b> ." | China, Guangdong, Hong Kong, Jiangmen, Taiwan                                                    | Sydney, Australia (33.874154, 151.180554)                                                         | Sydney, Glebe, Malaysia, Singapore | migration histories<br>community interactions |
| (2) "I'm in <b>Chicago Chinatown</b> . This is a replica of a Qing wall in Beihai Park in <b>Beijing</b> (which I saw when I was there in 2003)."                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Beijing                                                                                          | no geotag provided                                                                                | Chicago                            | aesthetic and term relations                  |
| (3) "The Khoo Kongsi clan house in the centre of <b>George Town, Penang, Malaysia</b> . Completed in 1906, the ornate clan house was the focal point of trade and community for <b>Hokkien Chinese</b> settlers. The building is beautiful both inside and out, and forms part of the city's UNESCO World Heritage site. Certainly a joy to walk around!"                                                                                                                                                                                                                                                                                                                                                                                         | Fujian, China                                                                                    | George Town, Malaysia (5.414566, 100.33733)                                                       | George Town, Penang, Malaysia      | migration histories<br>community interactions |
| (4) "I was entertained by this small alley, called <b>Tainan Alley</b> ."                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Tainan                                                                                           | Yokohama, Japan (35.4426, 139.645957)                                                             |                                    | aesthetic and term relations                  |
| Examples of Dutch dataset                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                  |                                                                                                   |                                    |                                               |
| (5) "Restaurant Sea Palace. The floating restaurant in the IJ River, was inspired by the floating restaurants of <b>Hong Kong</b> ."                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Hong Kong                                                                                        | Amsterdam, Netherlands (52.375172, 4.906317)                                                      |                                    | aesthetic and term relations                  |
| (6) "I mentioned before that because the Netherlands had close ties with <b>Indonesia</b> for several centuries, and ethnic <b>Chinese</b> have lived in <b>Indonesia</b> for over a century, there are <b>Indonesian-Chinese</b> restaurants in every town according to my <b>Dutch</b> friends. Hilvarenbeek is no exception. Plus there are also <b>Chinese</b> people who went to <b>Suriname/ Surinam</b> first who eventually moved to the <b>Netherlands</b> , and then there are <b>Hong Kong Chinese</b> who emigrated to <b>Holland</b> in the 1960s and 1970s, so there are all sort of <b>Chinese</b> restaurants all over the little country. Gelderstraat 16, Hilvarenbeek."                                                        | China, Hong Kong                                                                                 | Hilvarenbeek, Netherlands (51.484317, 5.137161)                                                   | Indonesia, Suriname, Netherlands   | migration histories                           |
| (7) "A story about the first Chinese restaurant in Alkmaar. They married, had two children and then decided to return to <b>China</b> in 1948, where they settled in <b>Shanghai</b> ...in 1952 the family decides to return to the <b>Netherlands</b> ."                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | China, Shanghai                                                                                  | Alkmaar, Netherlands (52.630054, 4.7484)                                                          | Netherlands                        | migration histories                           |
| (8) "Clear signs the taste of the <b>Dutch</b> is changing. The classic <b>Chinese</b> restaurants a being replaced by the popular <b>Japanese</b> kitchen."                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | China                                                                                            | Delft, Netherlands (52.010005, 4.360844)                                                          | Netherlands, Japan                 | changes in time                               |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |  placenames |  ethnicities |                                    |                                               |

Figure 3: Examples of the classification of ODR places, respectively from the global and Dutch datasets

### 3.2 Semantic Network of Noun Entities

Heritage attributes (elements of cultural significance needing protection) and interrelations of Chinese diasporic heritage can be partially represented in the semantic networks. “Amsterdam”, “Rotterdam” and “Netherlands” are core nodes with high connectivity. Diverse themes are shown in different cities (**Figure 4**).

Amsterdam has the largest number of nodes but lower average weights, reflecting multiple dimensions of heritage. Nouns related to tangible attributes are found, such as the streets and areas “Zeedijk”, “Geldersekade”, “red light (district)”, and the temple “Fo Guang Shan”. By the placename “Nieuwmarkt”, Amsterdam Chinatown is closely linked to the UNESCO World Heritage Property<sup>6</sup>, where a more dense subnetwork shares attributes like “narrow houses”, “seventeen century”, and “charming canals”. Rotterdam has more nouns with intangible meanings, like “Kung Fu”, “Chinese Light Festival”, and “Lion Dancing”. The Hague has the sparsest network. More nouns are related to public organisations, like “The Hague City Hall”, “Atrium”, and traditional elements, like “Chinese gates” and “red lanterns” (**Figure 5**).

---

<sup>6</sup>“Seventeenth-century canal ring area of Amsterdam inside the Singelgracht”

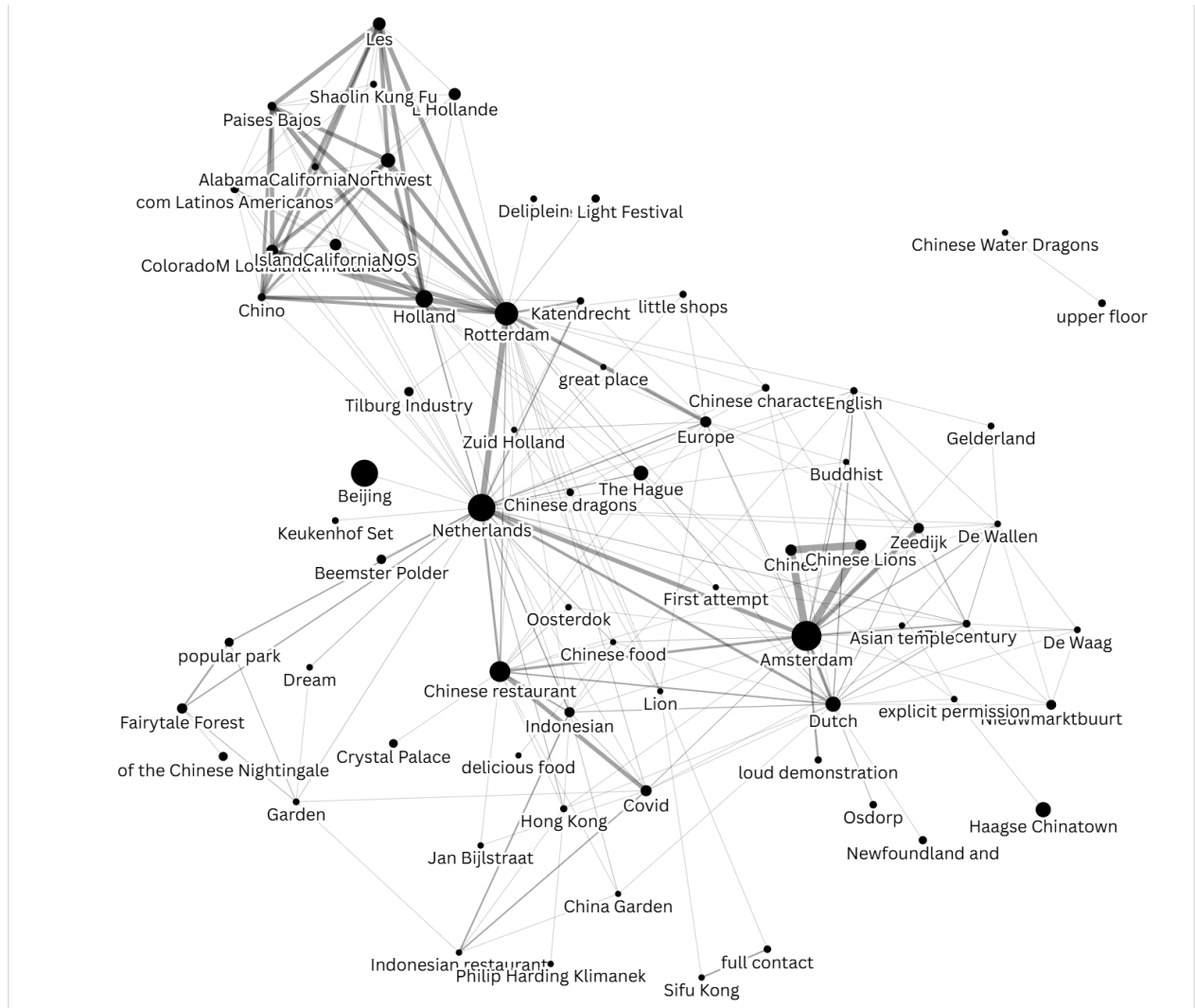
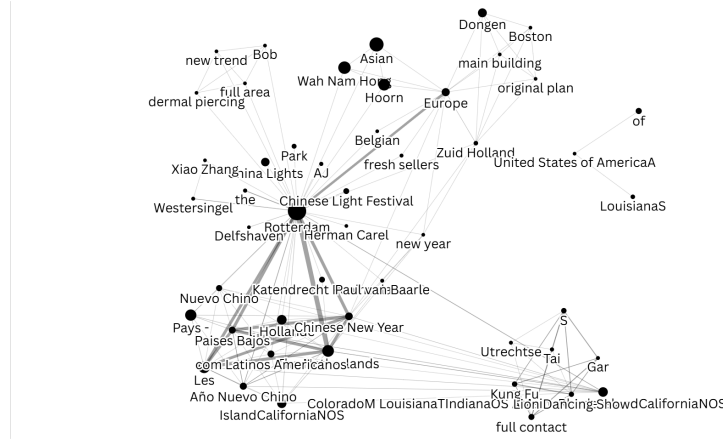


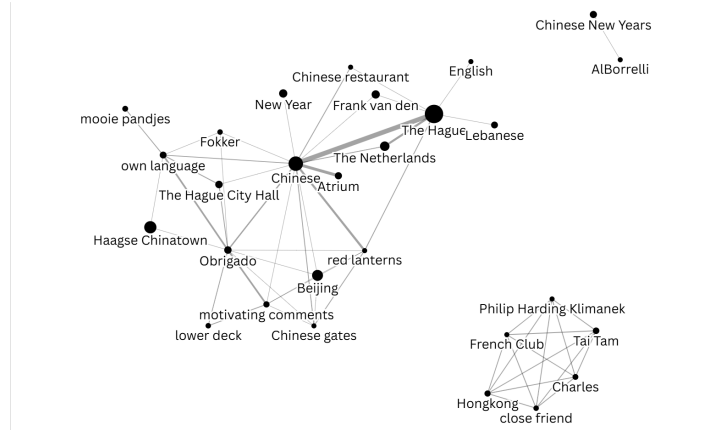
Figure 4: Semantic network of Chinese diasporic heritage in the Netherlands. Nodes with weights above 13 are displayed.



(a) Network of Amsterdam



(b) Network of Rotterdam



(c) Network of the Hague

Figure 5: Semantic networks of Chinese diasporic heritage comparing three cities in the Netherlands. Nodes with weights above 5 are displayed.

## 4 Discussion and Conclusions

This study provides a workflow for searching, mapping, and exploring diasporic heritage. The ODR network potentially represents flows of communities and heritage attributes. This attempt is promising for understanding the interactions between places of ODR and their impacts on local sites, challenging the territorialized stereotype of assessing cultural significance. The semantic network analysis distinguishes subnetworks in different cities, allowing the understanding of relationships between heritage attributes and places.

While social media data carries inherent risks of subjectivity, User-specific limitation, and noise, this workflow can be enhanced by engaging traditional qualitative methods and optimising data processing techniques. Despite these challenges, this study highlights the value of cross-geographical interrelations and social media data towards more comprehensive heritage identification, and more democratic urban planning and management.

This workflow can be used to analyse other diasporic communities or heritage shared on social media<sup>7</sup>. According to the spatio-temporal characteristics of ODR networks, future research can distinguish subgroups of diasporic communities, and determine different layers of heritage significance. Further statistical analysis and classification based on the semantic networks can also serve as supplementary data for heritage value assessment.

(word count: 1500)

## 5 Acknowledgements

This work was supported by China Scholarship Council [grant number 202306210072].

## References

- Shuval, Judith T. (2000). "Diaspora Migration: Definitional Ambiguities and a Theoretical Paradigm". In: *International Migration* 38.5, pp. 41–56. DOI: <https://doi.org/10.1111/1468-2435.00127>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-2435.00127>.
- Castles, Stephen (2007). "Twenty-First-Century Migration as a Challenge to Sociology". In: *Journal of Ethnic and Migration Studies* 33.3, pp. 351–371. ISSN: 1369-183X 1469-9451. DOI: 10.1080/13691830701234491.
- Lucassen, Jan and Leo Lucassen (2009). "The mobility transition revisited, 1500–1900: what the case of Europe can offer to global history". In: *Journal of Global History* 4.3, pp. 347–377. ISSN: 1740-0228 1740-0236. DOI: 10.1017/s174002280999012x.
- Fincher, Ruth et al. (2014). "Planning in the multicultural city: Celebrating diversity or reinforcing difference?" In: *Progress in Planning* 92. Planning in the multicultural city: Celebrating diversity or reinforcing difference?, pp. 1–55. ISSN: 0305-9006. DOI: <https://doi.org/10.1016/j.progress.2013.04.001>.

---

<sup>7</sup>As found on Flickr, there are groups such as "Tamil Diaspora Photography Arts Collective", "Little India - Singapore", "Asians of the World", "Pinoy Abroad!", etc.

- Byrne, Denis (2016). “Heritage corridors: transnational flows and the built environment of migration”. In: *Journal of Ethnic and Migration Studies* 42.14, pp. 2360–2378. ISSN: 1369-183X 1469-9451. DOI: 10.1080/1369183x.2016.1205805.
- Ginzarly, Manal, Ana Pereira Roders, and Jacques Teller (2019). “Mapping historic urban landscape values through social media”. In: *Journal of Cultural Heritage* 36, pp. 1–11. ISSN: 12962074. DOI: 10.1016/j.culher.2018.10.002.
- Lai, Juntao (2019). “Urban Place Profiling Using Geo-Referenced Social Media Data”. Thesis.
- Dellios, A. and E. Henrich (2020). “Migratory pasts and heritage-making presents: Theory and practice”. In: *Migrant, Multicultural and Diasporic Heritage: Beyond and Between Borders*. Ed. by Alexandra Dellios and Eureka Henrich. taylorfrancis, pp. 1–17. DOI: 10.4324/9780429328404-1.
- Danchev, Valentin and Mason A. Porter (2021). “Migration Networks: Applications of Network Analysis to Macroscale Migration Patterns”. In: *Research handbook on international migration and digital technology*. elgaronline.
- Ginzarly, Manal and F. Jordan Srouf (2021). “Unveiling children’s perceptions of World Heritage Sites: a visual and qualitative approach”. In: *International Journal of Heritage Studies* 27.12, pp. 1324–1342. ISSN: 1352-7258 1470-3610. DOI: 10.1080/13527258.2021.1977374.
- Bai, N. et al. (2023). “Mapping Urban Heritage Images with Social Media Data and Artificial Intelligence, a Case Study in Testaccio, Rome”. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII-M-2-2023, pp. 139–146. ISSN: 2194-9034. DOI: 10.5194/isprs-archives-XLVIII-M-2-2023-139-2023.
- Yuhang Gu Menno-Jan Kraak, Yuri Engelhardt and Franz-Benjamin Mocnik (2023). “A classification scheme for static origin–destination data visualizations”. In: *International Journal of Geographical Information Science* 37.9, pp. 1970–1997. DOI: 10.1080/13658816.2023.2234001.

## Biographies

Yan Zhou is currently a Ph.D. candidate at Faculty of Architecture and Built Environment, Delft University of Technology. She has her master’s degree in architecture history and theory at School of Architecture, Tsinghua University. Her research interests include Chinese architectural history, diasporic heritage, and participatory heritage management.

Nan Bai is Assistant Professor in section Heritage & Architecture at Delft University and Technology. With a multi-disciplinary background, Nan always has a broad interest in the application of data science, artificial intelligence, and computational social science in the historic urban settings, especially the perceptions reflected from social media.

Lidwine Spoormans is an Assistant Professor at the Faculty of Architecture and Built Environment at Delft University of Technology. She was trained at Delft University of Technology, worked in architecture practice and specialised in housing renovation, building systems, intervention methods, young heritage. Her research focuses on bridging the gap between everyday architecture and heritage.

Ana Pereira Roders is Professor in Heritage and Values, UNESCO chair on Heritage and the Reshaping of Urban Conservation for Sustainability, at TU Delft. Ana graduated as an Architect at University Lusíada; and obtained a PhD in Building Technology at TU/e. Her research focuses on

further understanding the values that define heritage, and how they impact the sustainability of cities.