


# Understanding Place Identity with Generative AI

**Kee Moon Jang** ✉ 

MIT Senseable City Lab, Cambridge, MA, USA

**Junda Chen** ✉


DataChat, Madison, WI, USA

**Yuhao Kang**<sup>1</sup> ✉ 

MIT Senseable City Lab, Cambridge, MA, USA

**Junghwan Kim** ✉ 

Department of Geography, Virginia Tech, Blacksburg, VA, USA

**Jinhyung Lee** ✉ 

Department of Geography and Environment, Western University, London, Canada

**Fábio Duarte** ✉ 

MIT Senseable City Lab, Cambridge, MA, USA

---

## Abstract

Researchers are constantly leveraging new forms of data to understand how people perceive the built environment and the collective place identity of cities. Latest advancements in generative artificial intelligence (AI) models have enabled the creation of realistic representations of real-world settings. In this study, we explore the potential of generative AI as the source of textual and visual information in capturing the place identity of cities assessed by filtered descriptions and images. We asked questions on the place identity of a set of 31 global cities to two generative AI models, ChatGPT and DALL·E2. Since generative AI has raised ethical concerns regarding its trustworthiness, we performed cross-validation to examine whether the results show similar patterns to real urban settings. In particular, we compared the outputs with Wikipedia data for text and images searched from Google for images. Our results indicate that generative AI models have the potential to capture the collective features of cities that can make them distinguishable. This study is among the first attempts to explore the capabilities of generative AI in understanding human perceptions of the built environment. It contributes to urban design literature by discussing future research opportunities and potential limitations.

**2012 ACM Subject Classification** Social and professional topics → Geographic characteristics

**Keywords and phrases** ChatGPT, DALL·E2, place identity, generative artificial intelligence, sense of place

**Digital Object Identifier** 10.4230/LIPIcs.GIScience.2023.41

**Category** Short Paper

**Related Version** *Previous Version*: <https://arxiv.org/abs/2306.04662>

## 1 Introduction

Place identity, often referred to as properties that distinguish a place from others [12, 11], is an important concept in the fields of urban design, geography, tourism, and environmental psychology. As a sense of place that is shaped through diverse human experiences, recognizing such place characteristics has been crucial for understanding human-environment interactions [4, 9, 10]. Yet, measuring and representing place identity has been a challenging task due

---

<sup>1</sup> Corresponding author: yuhaokang@mit.edu



to the intrinsically subjective nature of place identity. Conventional studies attempted to capture place identity through direct observation, questionnaires, surveys and interviews [4, 10]. In the past decade, researchers have been leveraging new data sources to understand the collective place identity of cities. In particular, two data formats, texts [4, 2, 3], and images such as street-level images and geotagged photos [16, 17] have been effective in revealing place identity information. Urban planners and designers have benefited from these emerging data sources to explore subjective urban experiences and promote data-driven decision-making processes in practices [10].

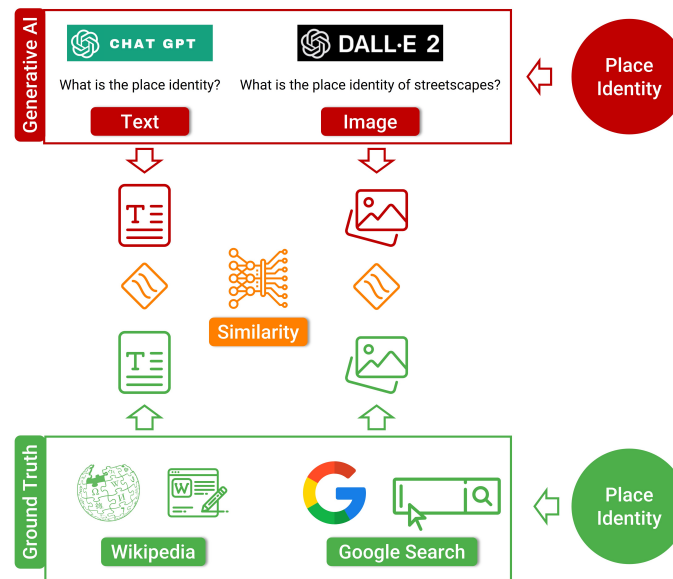
Recently, advancements in generative artificial intelligence (GenAI) models have received significant attention due to their capabilities to generate realistic text and image output based on natural language prompts. ChatGPT and DALL·E2, for instance, have been highlighted as powerful tools with the potential for a wide variety of applications in different domains such as education, transportation, geography, and so forth [6, 7, 8, 14]. Also, there have been attempts in urban studies to evaluate design qualities of the built environment scenes and obtain optimal land-use configuration through automated urban planning process [13, 15]. Despite its promise in urban science, the use of generative AI also faces common ethical concerns such as misinformation and bias, falling short in depicting composition and locales for specific conditions [5]. Therefore, there remains needs for a more robust quantitative examination and analysis of how well they represent place-specific contexts toward trustworthy outputs in different domains.

To this end, since generative AI models are offering new ways to collect textual and visual information that may represent realistic human responses, this study aims to examine the potential of generative AI as new tools for understanding place identity in different cities. In this endeavor, we address two research questions: (1) Can generative AI models identify place identity of cities? and (2) How reliable are the generated outputs when compared with real-world settings? This study is expected to guide urban researchers in using such tools to generate large volumes of data through a more efficient and cost-effective approach, as well as to study place identity in a data-driven manner, which can facilitate our understanding of urban perception.

## 2 Methodology

We present a computational framework of this study in Figure 1. The framework involves two datasets that we created to investigate the potential of generative AI models in capturing the place identity of 31 global cities. The first dataset is a text-based dataset that we generated using ChatGPT to understand place identity, using the following prompts: “*What is the place identity of {city}? Give me in ten bullet points*”. To ensure consistency and comparability across different cities included in our dataset, we limited the responses to ten bullet points. By doing so, the generated outputs are concise and structured, and can easily be analyzed and compared. The second dataset is an image-based dataset that we collected using DALL·E2 to generate visual representations of streetscapes in different cities. The prompts used to achieve this are the following: “*What is the place identity of streetscapes of {city}*”? We generated 10 images for every city, where each image has a size of 256\*256 pixels. By combining the image-based dataset with the text-based dataset, we aim to provide a comprehensive and multi-modal understanding of the place identity of each city.

We further collected two ground-truth datasets including a text dataset from Wikipedia and an image dataset from Google search. Despite the high performance of generative AI tools in generating realistic outputs, concerns regarding their reliability and accuracy have emerged. Thus, we performed the cross-validation to compare the similarities among



■ **Figure 1** The computational framework of this paper.

these datasets to evaluate whether the results provided by generative AI can be trustworthy. For text similarity, we first segmented the Wikipedia corpus into individual sentences and converted each sentence from both datasets into word embeddings. This was achieved by using a sentence transformer BERT model based on a modified version of MiniLM. Then, we measured cosine similarity for sentence embeddings from ChatGPT responses and Wikipedia corpuses to assess the relevance between the two datasets. We also created word cloud images of each city for a visual comparison between topics covered in ChatGPT and Wikipedia texts.

For image similarity, we measured the Learned Perceptual Image Patch Similarity (LPIPS) [18] to assess the perceptual similarity of images generated by DALL · E2 and collected via Google search. The LPIPS metric evaluates the distance between different image patches and produces scores ranging from 0 to 1, where a lower score indicates greater similarity, and vice versa. Subsequently, we identify the top 3 similar Google images for each DALL · E2-generated image based on similarity scores. These analyses allow us to validate whether the results generated by generative AI models are consistent with the real-world urban settings of each city, providing valuable insights for urban design research and practice.

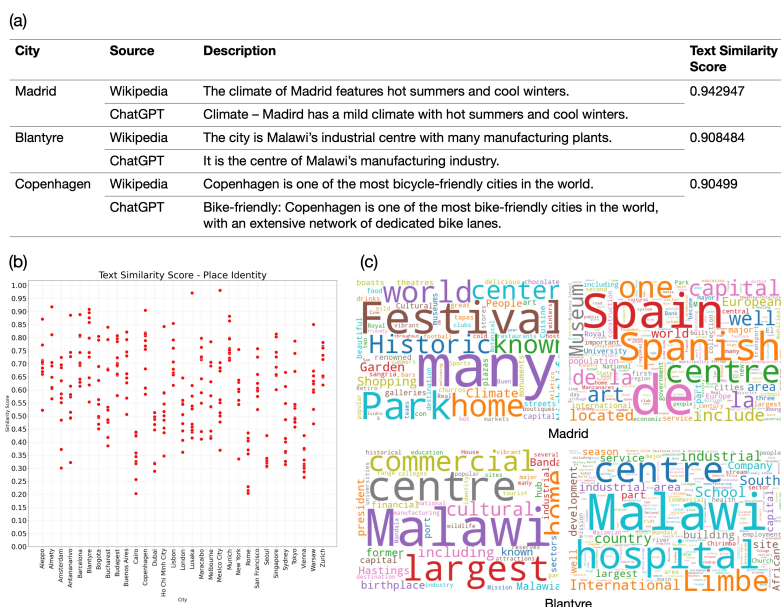
## 3 Results

### 3.1 Results of place identity generated by ChatGPT

To validate the accuracy and reliability of the data generated by ChatGPT, we conducted cross-validation with Wikipedia. This involved computing the similarity scores between sentences from ChatGPT and Wikipedia, and presenting visual comparisons between pairs of word clouds. Figure 2 illustrates the validation results. Figure 2(a) shows several examples of high sentence similarity scores. For example, for the city of Madrid, both Wikipedia and ChatGPT-generated sentences had similar descriptions of the climate, resulting in a very high similarity score of 0.94. However, as shown in Figure 2(b), the comparison between

## 41:4 Understanding Place Identity with Generative AI

ChatGPT-generated sentences and the introduction from Wikipedia resulted in a range of similarity scores, reflecting both similar and dissimilar descriptions of place identity. Such disparities may suggest that there are limitations to the effectiveness of generative AI models in capturing the nuances and complexities of place identity. Last, Figure 2(c) illustrates two cases of word clouds analysis created for ChatGPT responses (left) and Wikipedia (right). While the introduction of Madrid in Wikipedia covered generic keywords such as *Spain*, *Spanish*, *centre* and *capital*, we found that ChatGPT captures the full spectrum of place identity components as defined in fields of environmental psychology and geography [1, 12]. For instance, topics including *park* and *garden* refer to the “physical settings” of Madrid, *festival* and *shopping* represent the “activities” that take place, and *historic* and *climate* describe the subjective “meanings” that contribute to the place identity formation in the capital of Spain. In the case of Blantyre, the most notable keywords observed in the word cloud of Wikipedia corpus are *Malawi* and *centre*. Likewise, the word cloud generated from ChatGPT response also features the same keywords, from which we infer that ChatGPT identifies the place identity of Blantyre in relation to its significance within the national context of Malawi.

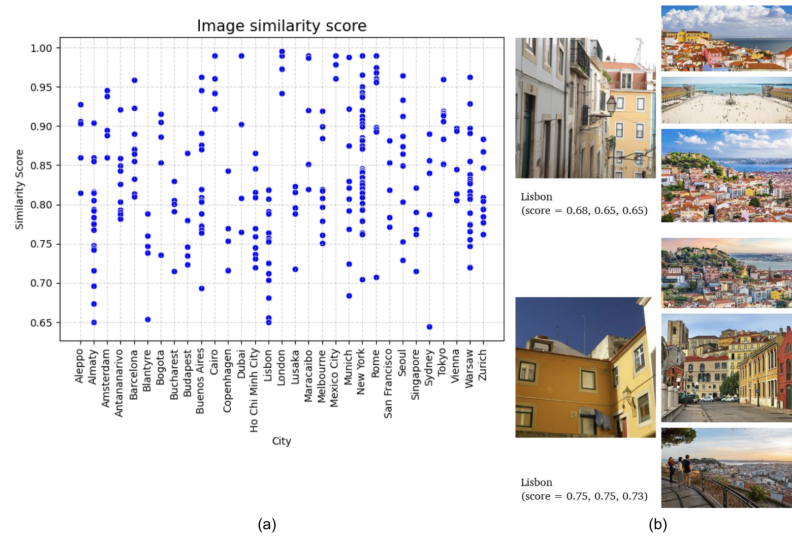


■ **Figure 2** Text similarity results. (a) Examples of high text similarity scores between Wikipedia introductions and ChatGPT responses on place identity; (b) Scatter chart of the distribution of cosine similarity scores between sentences from ChatGPT responses and Wikipedia introduction corpuses; and (c) Word cloud comparison for Madrid and Blantyre cases.

### 3.2 Results of place and urban identity generated by DALL · E2

Similar to the comparison between ChatGPT-generated sentences with Wikipedia corpus, we also compared images generated by DALL · E2 and those collected from Google search. This was conducted to verify the reliability and generative capability of the text-to-image model in producing realistic representations of place-specific scenes of cities. For this purpose, the image similarity was measured using LPIPS metric to assess the perceptual similarity between AI-generated and real-world images that match well with human judgment. Figure 3

presents the image similarity results. Overall, as shown in Figure 3(a), Almaty, Blantyre, Lisbon and Sydney were cities that reported the highest perceptual similarity with LPIPS value being approximately 0.65. In particular, Lisbon presents relatively consistent low similarity scores within the range of 0.65-0.82. Figure 3(b) shows two examples of DALL · E2 generated images for Lisbon's place identity and their top three matching Google image search results. It is evident that the generative AI effectively captured the low-rise residential buildings with vivid yellow colors in Lisbon, resulting in a low LPIPS score (high similarity). These suggest that, despite variability across cities, DALL · E2 can generate more reliable images of urban scenes for certain cities that reflect their place identity.



■ **Figure 3** Image similarity results. (a) Distribution of LPIPS scores between DALL · E2 generated images and Google image by cities; and (b) Low LPIPS score examples for Lisbon case.

## 4 Conclusion

In this study, we presented text and image similarity results between responses from two generative AI model, ChatGPT and DALL · E2, and corresponding ground-truth data to test the reliability of their outputs for representing place identity of different cities. Through examining the two datasets, we find that, in many cases, they generated text description or realistic images that represent salient characteristics of cities. In particular, text similarity scores aligned closely with similarities observed in sentence-by-sentence comparison and word clouds of ChatGPT responses and Wikipedia corpuses. This study is among the first to examine the capabilities of generative AI tools in representing the place identity of cities. The overall framework is expected to aid planners and designers in utilizing such tools to identify salient characteristics of cities for sustainable placemaking and city branding purposes.

Despite the contributions of this study, we discuss potential limitations and research opportunities to be addressed in future studies. First, a portion of DALL · E2 generated images is still considered more generic than place-specific, which may not fully reflect the place identity. These images are more relevant to the generic concept of a *city*, rather than *identity*, and fall short in representing the attributes that distinguish a particular city from the rest. Another limitation lies in the uncertainty in the image similarity results. We found that certain similar scenes generated by DALL · E2 resulted in a range of similarity

scores when measured against the same ground-truth image. Yet, it is uncertain why such differences are observed, what contributes to high or low similarity results, and thus which scene is most relevant to the place identity of a particular city.

---

## References

- 1 David Canter. *The psychology of place*. St Martin'S Press, 1977.
- 2 Song Gao, Krzysztof Janowicz, Daniel R Montello, Yingjie Hu, Jiue-An Yang, Grant McKenzie, Yiting Ju, Li Gong, Benjamin Adams, and Bo Yan. A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science*, 31(6):1245–1271, 2017.
- 3 Yingjie Hu, Chengbin Deng, and Zhou Zhou. A semantic and sentiment analysis on on-line neighborhood reviews for understanding the perceptions of people toward their living environments. *Annals of the American Association of Geographers*, 109(4):1052–1073, 2019.
- 4 Kee Moon Jang and Youngchul Kim. Crowd-sourced cognitive mapping: A new way of displaying people's cognitive perception of urban space. *Plos one*, 14(6):e0218590, 2019.
- 5 Yuhao Kang, Qianheng Zhang, and Robert Roth. The ethics of ai-generated maps: A study of dalle 2 and implications for cartography. *arXiv preprint arXiv:2304.10743*, 2023.
- 6 Junghwan Kim and Jinhyung Lee. How does chatgpt introduce transport problems and solutions in north america? *Findings*, 2023.
- 7 Ehsan Latif, Gengchen Mai, Matthew Nyaaba, Xuansheng Wu, Ninghao Liu, Guoyu Lu, Sheng Li, Tianming Liu, and Xiaoming Zhai. Artificial general intelligence (agi) for education. *arXiv preprint arXiv:2304.12479*, 2023.
- 8 Gengchen Mai, Weiming Huang, Jin Sun, Suhang Song, Deepak Mishra, Ninghao Liu, Song Gao, Tianming Liu, Gao Cong, Yingjie Hu, et al. On the opportunities and challenges of foundation models for geospatial artificial intelligence. *arXiv preprint arXiv:2304.06798*, 2023.
- 9 Lynne C Manzo and Douglas D Perkins. Finding common ground: The importance of place attachment to community participation and planning. *Journal of planning literature*, 20(4):335–350, 2006.
- 10 Mahbubur Meenar, Nader Afzalan, and Amir Hajrasouliha. Analyzing lynch's city imageability in the digital age. *Journal of Planning Education and Research*, 42(4):611–623, 2022.
- 11 Harold H Proshansky, Abbe K Fabian, and Robert Kaminoff. Place-identity: Physical world socialization of the self (1983). In *The people, place, and space reader*, pages 111–115. Routledge, 2014.
- 12 Edward Relph. *Place and placelessness*, volume 67. Pion London, 1976.
- 13 Sachith Seneviratne, Damith Senanayake, Sanka Rasnayaka, Rajith Vidanaarachchi, and Jason Thompson. Dalle-urban: Capturing the urban design expertise of large text to image transformers. *arXiv preprint arXiv:2208.04139*, 2022.
- 14 Eva AM van Dis, Johan Bollen, Willem Zuidema, Robert van Rooij, and Claudi L Bockting. Chatgpt: five priorities for research. *Nature*, 614(7947):224–226, 2023.
- 15 Dongjie Wang, Chang-Tien Lu, and Yanjie Fu. Towards automated urban planning: When generative and chatgpt-like ai meets urban planning. *arXiv preprint arXiv:2304.03892*, 2023.
- 16 Fan Zhang, Bolei Zhou, Carlo Ratti, and Yu Liu. Discovering place-informative scenes and objects using social media photos. *Royal Society open science*, 6(3):181375, 2019.
- 17 Fan Zhang, Jinyan Zu, Mingyuan Hu, Di Zhu, Yuhao Kang, Song Gao, Yi Zhang, and Zhou Huang. Uncovering inconspicuous places using social media check-ins and street view images. *Computers, Environment and Urban Systems*, 81:101478, 2020.
- 18 Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018.