GOOD SYSTEMS
Ethical AI at UT Austin

The University of Texas at Austin
Research Development
Vice President for Research, Scholarship and Creative Endeavors

Inside Out: GeoAI and Human Perspective on the Quality of Urban Environments

Ahyoung Chang[1], Junfeng Jiao[1], Gengchen Mai[2]
[1]Urban Information Lab, School of Architecture
[2]Department of Geography and the Environment, The University of Texas

Introduction

Urban experience flows across both outdoor public spaces and indoor environments, yet conventional urban research has focused mainly on the external city. Using Generative AI, this study evaluates the Quality of Place (QoP) across ten global cities by analyzing Google Street View images (outdoors), Airbnb listing images (indoors), and guest reviews (human perceptions). Outdoor spaces were interpreted through Kevin Lynch's urban form elements, while indoor spaces were assessed using Vitruvius's triad, and guest reviews were processed with NLP to capture experiential and emotional meaning.
Comparing AI-generated descriptions with human perceptions reveals clear asymmetries. In cities like Amsterdam, Paris, and Sydney, AI and humans similarly emphasized coherence and walkability; however, in Hong Kong, Mexico City, and New York, AI prioritized visual order while overlooking sensory, cultural, and relational qualities noted in reviews. These differences highlight how vision-centered AI frameworks can reproduce spatial bias, underscoring the need to incorporate lived experience and cultural context into GeoAI-based urban evaluation.

Research Goal

- **Integrate** indoor and outdoor environments in evaluating Quality of Place (QoP)
- **Generate** spatial descriptions using Generative AI (image → text)
- **Extract** human perceptions from user-generated reviews (text → sentiment/meaning)
- **Compare** AI-based and human-based spatial evaluations across global cities

Methods

For each city, we collected three types of data: outdoor environments using Google Street View images, indoor environments using Airbnb listing images, and human perception data from Airbnb guest reviews. The outdoor images were evaluated with a **Generative AI prompt** framework based on Kevin Lynch's five elements of urban form (paths, edges, districts, nodes, and landmarks). Indoor images were analyzed using a parallel prompt structure grounded in Vitruvius's triad (firmitas, utilitas, venustas). Guest reviews were processed with **natural language processing** to extract sentiment and spatial meaning, capturing how visitors describe comfort, neighborhood accessibility, and everyday lived experience. Finally, we compared AI-generated spatial evaluations with human perceptions using **topic modeling** and **cosine similarity** analysis to examine where machine interpretations align with or diverge from experiential narratives.

GenAI-based Quality of Place Analysis

Research Scope:

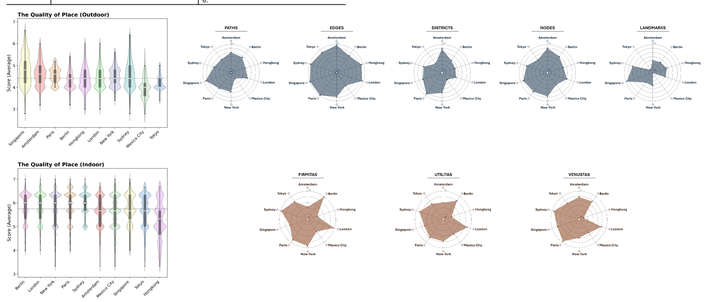
Amsterdam, Berlin, Hongkong, London, Mexico City, Sydney, New York, Tokyo, Paris, Singapore

Elements	Amsterdam – High (Score: 6.2)	Amsterdam – Low (Score: 1.8)
Paths	The tram tracks and **paved walkways/roads are visually clear** and appear easily navigable for pedestrians and cyclists.	The paths appear **discontinuous** and unclear, with a mix of paved and **unpaved** areas, making navigation seemingly difficult.
Edges	The buildings along the street create a clear edge, defining the **open square** from the built environment, while the large wooden wall acts as a temporary but prominent edge.	The edges are **poorly defined**, with a mix of vegetation, fences, and structures providing weak boundaries.
Districts	The area appears to be a **mix of commercial and public space**, with retail stores along the edge and a large open square serving as a gathering place, creating a coherent urban scene.	The area **lacks a clear distinctive character** or function, appearing as a transitional zone between different land uses with low visual coherence.
Nodes	The open square functions as a **central** node, where paths converge and people gather, providing a **focal point for activity.**	There are no clear focal points or areas where paths converge or people might gather, indicating a **lack of social or functional nodes.**
Landmarks	The large building in the background (likely the Royal Palace) serves as a prominent landmark, contributing to the city's identity and providing **orientation** within the urban space.	The image **lacks any visually** striking or culturally significant features that could serve as landmarks, offering little to contribute to a sense of place.

Elements	Tokyo – High (Score: 6.2)	Tokyo– Low (Score: 2.4)
Paths	The paths, including roads and crosswalks, are visually **clear, continuous,** and appear **highly navigable.**	The narrow path between the **construction walls** appears navigable but uninteresting and uninviting.
Edges	The edges are defined by roads, sidewalks, and changes in elevation, creating distinct zones between the park, road, and buildings.	The large, continuous construction walls create a strong, but monotonous and unattractive, edge, separating the path from the surrounding urban context.
Districts	The image suggests a mix of districts, potentially including **residential, commercial, and park areas**, though the distinctions aren't entirely clear from this single perspective.	The area within the construction walls lacks a distinctive character, appearing as a transitional zone between potentially more defined districts outside of the image frame.
Nodes	The intersection serves as a clear node where paths converge, and the open area near the park suggests a potential gathering space.	There are **no discernible nodes** or points of intersection within the provided view, suggesting a lack of gathering spaces or significant activity within this particular area.
Landmarks	The **Tokyo Tower** serves as a prominent landmark, providing a strong sense of identity and orientation.	The image lacks distinct landmarks, with the surrounding buildings barely visible and not contributing to a clear identity or sense of place.

Elements	Amsterdam – High (Score: 6.7)	Amsterdam – Low (Score: 3.3)
Firmitas	The solid walls, countertops, and integrated appliances suggest a robust and stable structure.	The space appears structurally sound with well-defined walls and ceiling, although the exposed beams and pipes suggest potential for improvement in concealing infrastructure.
Utilitas	The open-plan layout combines kitchen and dining functions with ample light and accessibility, although the limited view of other areas makes it difficult to fully assess privacy and other potential uses.	The layout seems somewhat cramped and awkwardly arranged, with limited privacy between the kitchen/hallway area and the visible bedroom, and the luggage rack placement suggesting a lack of dedicated storage space.
Venustas	The neutral color palette, natural materials, and balanced composition of elements like the shelves, flowers, and lighting create a visually harmonious and aesthetically pleasing space.	The space lacks visual appeal due to the clashing styles, including the dated Coca-Cola advertisement, the basic kitchen appliances, and the mismatched rug and floor tiles, resulting in a disharmonious atmosphere.

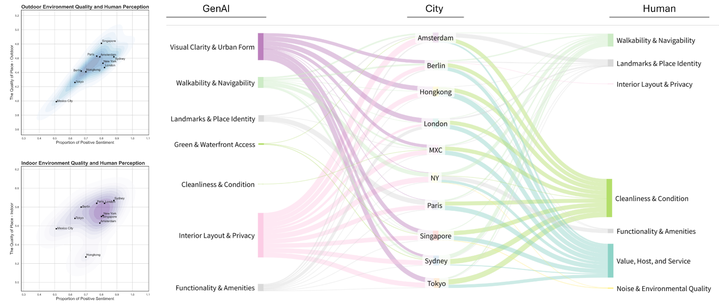
Elements	Tokyo – High (Score: 6.7)	Tokyo– Low (Score: 3.0)
Firmitas	The concrete structure, glass partitions, and stone flooring appear robust and well-integrated, suggesting excellent structural stability.	The room appears structurally sound with well-defined walls and ceiling, though the unusual orientation of the furniture suggests potential instability or unconventional design choices.
Utilitas	The layout seems to balance openness and privacy, with defined areas for relaxation and potentially other functions, though the exact purpose isn't entirely clear from the image alone.	The layout seems cramped and poorly optimized for functionality, with furniture awkwardly placed and a lack of clear division between different areas.
Venustas	The space achieves a harmonious aesthetic through the interplay of natural materials, warm lighting, and thoughtful design elements like the water feature and curated furniture, creating a visually appealing and tranquil atmosphere.	The room lacks visual harmony and aesthetic appeal, with mismatched furniture, cluttered surfaces, and an overall disorganized appearance.



Singapore, Sydney, and Amsterdam consistently scored high in outdoor QoP—particularly in paths and edges—indicating walkable and visually coherent urban form, while **Mexico City and Hong Kong** showed lower outdoor cohesion. Indoor evaluations showed a different geography, with Amsterdam, Tokyo, and Berlin ranking highest, and New York and Mexico City lagging in functional utility.

When comparing AI-evaluated QoP with human sentiment from reviews, we observed an overall positive correlation, yet with meaningful divergences: cities like Singapore and Amsterdam aligned strongly between AI perception and lived experience, while Hong Kong and Berlin showed moderate AI scores but lower sentiment, suggesting sensory or experiential discomfort not captured visually.

AI vs Human



Topic modeling further confirmed this gap—cities with legible street structure and coherent spatial identity (e.g., Amsterdam, Sydney, Paris) exhibited high semantic similarity between AI and human narratives, whereas Hong Kong and Mexico City revealed mismatches, with AI emphasizing visual form and humans emphasizing noise, comfort, service, and cultural atmosphere. These results highlight both the potential and the limitations of vision-centered GeoAI, revealing where machine perception aligns with lived experience—and where embodied, cultural, and sensory dimensions remain unseen.

Conclusion

These findings demonstrate that while Generative AI can systematically evaluate spatial environments, its perception remains fundamentally visual. Humans experience cities as continuous, affective environments—shaped by movement, sound, comfort, safety, and social atmosphere—whereas AI interprets static visual cues and overlooks sensory and relational qualities. We also observe a geographic pattern in which European cities tend to score higher in AI evaluations, reflecting aesthetic biases embedded in training data. Therefore, AI-based QoP assessment should be understood as complementary rather than substitutive; it can reveal structural and visual patterns at scale, but it cannot replace lived experience. To support equitable and meaningful urban planning, GeoAI systems must be designed with awareness of cultural and sensory diversity, and incorporate human voices and participatory interpretation in the evaluation process.

Acknowledgement

This research was supported by the NSF Grants (2043060, 2133302, 1952193, 2125858, 2236305).