AGILE: GIScience Series, 5, 22, 2024. https://doi.org/10.5194/agile-giss-5-22-2024
Proceedings of the 27th AGILE Conference on Geographic Information Science, 4–7 Sept. 2024.
Editors: Alison Heppenstall, Mingshu Wang, Urska Demsar, Rob Lemmens, and Jing Yao.
This contribution underwent peer review based on a full paper submission.

Construction of a walkability quality dataset in GIS: the case study of Berlin

Federica Gaglione

University of Sannio, Department of Engineering, Benevento, Italy.

Correspondence: Federica Gaglione (fgaglione@unisannio.it)

Abstract. This work aims to define and build a pedestrian network database by exploiting the information and potential of Google Street View. Compared to existing data such as Open Street Map, this dataset offers more detailed information on the quality of the street from a walker's point of view, also including urban context and safety characteristics in addition to those related to the geometry of a route. The research work is experimentation on the Berlin case study. The construction of the database allows, in turn, to obtain thematic maps of network classification as a result, providing a first element for local decision makers on where to implement possible improving solutions.

Keywords. Urban Accessibility; GIS; network analysis

1 Introduction

In recent years, the theme of walkability has clashed with the difficulties of providing computational-scientific support in analyses, processing, and experiments in trying to provide possible cognitive scenarios in different territorial contexts.

Data-driven quantitative scientific research on the topic of walkability has been addressed according to two types of approaches (Yin, 2017; Li et al., 2023). The first of the "place-based" type, aimed at measuring pedestrian levels for minimum census sections or catchment areas due to the presence of local services (for example green areas, recreational activities, healthcare) (Boulange et al., 2017; Suminski et al., 2023). A second approach, however, of the "network-based" type aims to measure the accessibility to services at a local scale by focusing on the characteristics of the pedestrian network and on the identification of the best routes for users (Su et al., 2019; S. Alfosool et al.,2022). There are few studies that consider network characteristics linked to connectivity, land use and the presence of services without considering specific network characteristics linked for example to the presence of traffic lights as well as the state of the pavement and therefore linked more to the quality of the route that the user takes in the city. The few studies that take these characteristics into account are aimed at a smaller scale of the neighbourhood (Loo et al., 2012). In turn, decision support tools such as Walk Score® (https://www.walkscore.com/) or Pedestrian Quality Needs' (https://pathways.walk21.com/dashboard/) which have European funding behind them or companies do not present specific details on walkability quality in their interactive maps. For example, the commercial web tool Walk Score® is primarily used by users to evaluate their decision to buy a house, and by researchers to measure characteristics of the built environment in transport and health research. Instead, the Walk 21 platform evaluates walking indicators for five macro categories which are: (i) policy; (ii) activity; (iii) safety; (iv) accessibility; (v) comfort on scales such as country or regional by estimating a partial and overall score on a radar diagram using a scale of values from 1 to 5.However, the measurement of the walkability of the built environment cannot be done without the most accurate level of detail compatible with the resources available. This paper presents a mapping exercise carried out with students at the University of Sannio, which resulted in the enrichment of Berlin's Street network data with information relevant to capture the quality of the walk. Assessing how walkable streets are is a fundamental task to inform policies aimed at supporting sustainable mobility in cities. However, a comprehensive assessment of the quality of streets for walking is often limited by the lack of detailed data. In the context of an Urban Planning Technique course, some students volunteered to add the missing information to assess the walking quality on the street network of Berlin districts.

Through GIS training, the support of Street view images and Google Traffic service, students with the supervision and support of the course organizer, were able to create a novel dataset for Berlin that can be used to develop more detailed walkability measures. The paper is structured as follows: section 2 presents a literature review on what makes streets walkable that supported the variable selection, section 3 describes the data collection process, section 4 shows the maps resulting from the data collection and section 4 highlights the possible applications of the dataset.

2 Defining a good quality walk

The methodological work of this research consisted of two parts. The first step involved identifying the set of variables useful in defining the quality of walking for a user. In this direction, a review of the relevant scientific literature was undertaken to define which variables are most significant in measuring a user's pedestrian accessibility at the urban scale (Koo et al.,2023; Valls & Clua, 2023). The scientific debate regarding the study of pedestrianism encompasses three main aspects. The first aspect focuses on investigating about the geometry of a path such as, for example, the width of the sidewalk or the connectivity of the pedestrian network (D'Orso & Migliore, 2020). Some studies have pointed out that high levels of connectivity indicate that there are many options for a pedestrian to choose different routes to travel between two specific points in an urban area (Chan & Farber, 2020; Liu et al.,2020). In contrast, Moon et al. (2016) pointed out that sidewalk quality (e.g., the type of material used for a sidewalk) and the absence of steps are among the main factors that encourage walking (Gaglione et al.,2021). The second aspect on which studies have focused on the perceived safety of the built environment and which can reduce the risk of pedestrians falling due to slippery conditions, the risk of conflicts between pedestrians and vehicles (Southworth, 2005). Studies emphasize the importance of certain pedestrian requirements such as traffic signals, the presence of crosswalks (Stipancic et al.,2020; Lee et al.,2019; Ruiz-Padillo et al.,2018). Finally, the third aspect related to the attractiveness and comfort of a pedestrian route. Studies have pointed out that the presence and quality of amenities (e.g., green areas) and shaded open spaces can favor the pleasantness and amenity of a pedestrian route over others (Li et al.,2023; Loo et al.,2012). In this direction, this study considers, in accordance with the scientific literature, a set of variables divided into three categories: (i) physical characteristics that refer to the geometry and quality of pedestrian routes; (ii) safety characteristics such that support, protection and sense of security are perceived toward places and services; and (iii) urban context characteristics that refer to the attractiveness of a pedestrian route. Table 1 makes explicit for each macro category the variables considered. Some variables were parameterized at the quantitative level and others at the qualitative level. Variables that were parameterized on a qualitative scale, such as the condition of the pavement, were defining evaluation criteria (D'Orso & Migliore, 2020). The pavement condition was considered poor in the presence of holes or dips, fair if the pavement is degraded and finally good in absence of holes or dips, pavement in a good state. The definition of the set of variables was the starting point for the construction of the data and database aimed at the case study of the city of Berlin, and which is still lacking in the scientific frame of reference and in the development of decision support tools. This methodological approach has already been developed and tested in other national and international contexts such as Naples and Aberdeen (Gaglione et al.,2021). Only that the research work elaborates a dataset for Naples and Aberdeen limited to a neighborhood scale compared to the case study of Berlin, whose work is at an urban scale.

Table 1. Set of pedestrian network characteristics variables

Variable	Unit of measure	References
Physical characteristics		
Network connectivity	n of intersections / link length	Chan & Farber, 2020; Liu et al.,2020
Sidewalk width	meters	Valls & Clua, 2023; D'Orso & Migliore, 2020
State of sidewalk pavement	Poor/fair/ good	Moon et al., 2016; Gaglione et al.,2021
Safety characteristics		
Crossroad	yes/no	Southworth,2005; Stipancic et al.,2020
Pedestrian crossings	yes/no	Stipancic et al.,2020; Ruiz-Padillo et al.,2018
Speed of vehicular traffic	km/h	Ruiz-Padillo et al.,2018; Stipancic et al.,2020
Presence of traffic lights	yes/no	Gaglione et al.,2021; D'Orso & Migliore, 2020
Urban context characteristics		
Presence of benches	yes/no	Li et al.,2023; Loo et al.,2012
Presence of shaded paths	yes/no	Li et al.,2023; Loo et al.,2012

3 Data Collection

The second part of this work was devoted to collecting data that enrich the description of the street network with information on its quality from the perspective of a pedestrian. Building upon the definition of what a good quality walk is, new data were collected for the city of Berlin. Berlin has a very complex structure, extending over a land area of approximately 891.12 km² divided into 12 districts. As of today, it has a population of approximately 3,664,088 with a population density of 4,117.8 Inhabitants/sqm. Its complexity and rapidly changing present, makes Berlin an interesting case study. The workflow of the data collection is presented in Fig. 1. The data was collected by civil engineering students Urban Planning Technique course (BSc Bachelor of Science) at Sannio University. The primary objective of the course was education on the GIS tool. Students in the preliminary phase acquired the GIS skills required to build the dataset. Each student was responsible for studying a particular district in the city of Berlin. Subsequently, meetings were held twice a week for two hours during the three-month duration of the course. This resulted in a crowd mapping exercise with students.

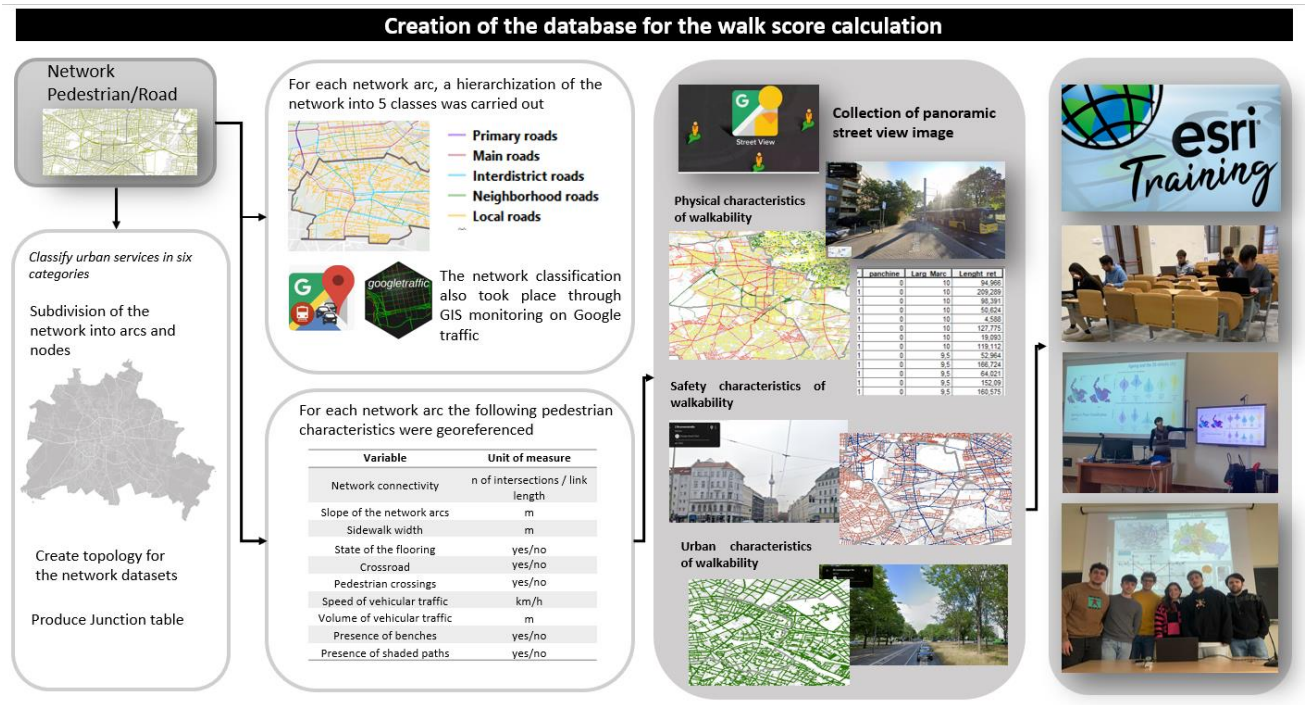

Figure 1. Workflow for the creation walk score.

The first step involved the hierarchization of the network divided into five classes: primary roads; principal; inter-district; neighborhood and local. In this research work, primary roads mean roads with their own entrances and exits and which ensure a level of long-distance movement. By principal roads, however, we mean the roads that connect each district with the rest of the territory and which at the same time play the role of distributing the main traffic flows. Inter-district roads are those roads that connect the different districts of the city. Neighborhood streets are the arcs of the network that provide the most important connections within the districts of the city itself. Finally, local roads directly serving buildings. Once the hierarchy of the road network had been defined, the maximum speed of each network arc was examined to understand which network arcs had a pedestrian-type reward as a basis. Furthermore, to capture the volume of vehicular traffic, the network classification was also accompanied by monitoring the traffic flows for each day of the week at 8.00, 12.00, 16.00 and 20.00 using the Google Maps Traffic application. The use of Street View Imagery (SVI) enabled a detailed assessment of physical walkability, safety, and urban context by extracting a variety of information in the walkability assessment. The students did not only examine individual urban characteristics, but through a process of normalization of the variables, the potential of walkability was also assessed for individual groups of characteristics that collectively.

Once the data collection and geo-referencing was complete, user behavior in terms of walking speed was investigated (Figure 2). Specifically, four populations were considered: (i) children; (ii) disabled; (iii) elderly; (iv) average users. For each group, travel times on the network were defined. These were then classified into five classes (5;10; 15; 20; > 20 min)

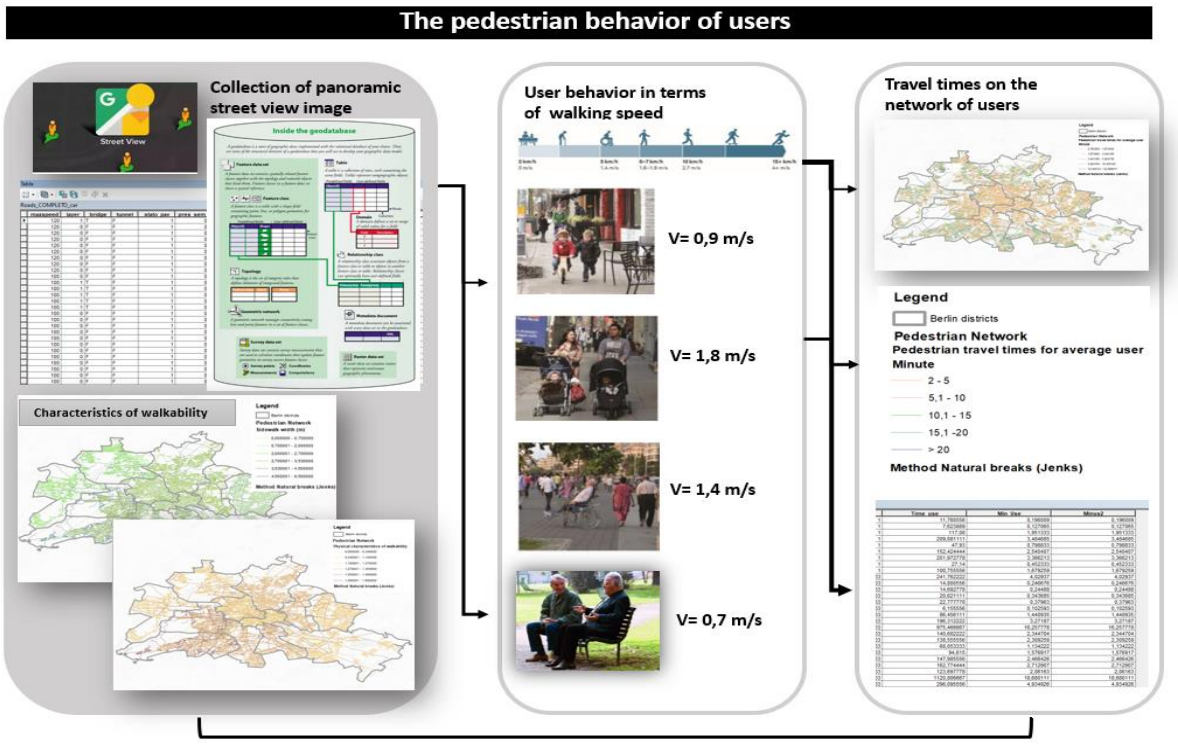

Figure 2. Data collection and the pedestrian behaviour of users

4 Mapping

The data collection phase resulted in a novel dataset for the city of Berlin providing more detailed information on the quality of the streets from the perspective of a walker.

Figure 3 shows the variables that were collected by the students in the Neukölln borough. This borough resulted as the most complete and rich of information after the mapping exercise.

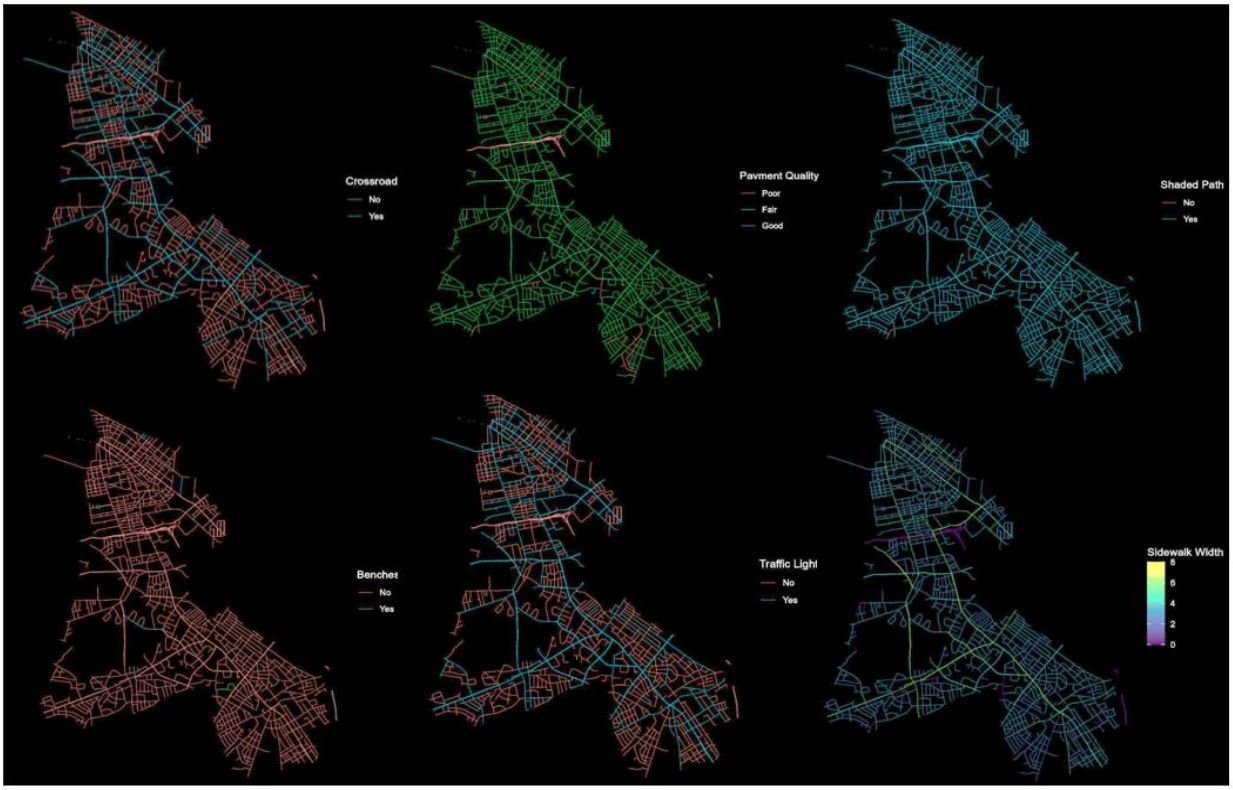

Figure 3. The quality of the streets of a walker.

It can be noted that most of the streets are shaded and have a pavement of good quality; the provision of safe crossings and traffic light is mostly concentrated in the main roads that are also equipped with wide sidewalk. There is a clear lack of benches provision, which was unexpectedly noted also in the more central boroughs.

5 Conclusions

This study highlights the value of information, and in particular geographical information, for disciplines characterized by a strong spatial connotation (Murgante et al., 2009). The study of urban phenomena is a scientific field that is largely based in the cognitive phase on the study of the "measurement" of physical entities such as the mobility system, of functional entities such as urban services and on the spatial analysis of such phenomena. This work was a first experiment engaging students in learning about walkability by crowd mapping information on the characteristics of streets that are key to make a walk more pleasant. Students, through their educational experience, were also able to generate a new dataset that can be used to develop new walkability indexes or generate alternative routes based on the quality of the street.

Acknowledgments

- I would like to thank the students Marco Iannelli, Giovanni Melillo, Stefano Meola, Siria Ruggiero, Stefano Terlizzi and Luca Zeoli who, as part of the Urban Planning Technique course at the University of Sannio, have contributed to supporting scientific research by constructing a dataset on the walkability of the city of Berlin as part of their training in the GIS tool.

References

Boulange, C., Pettit, C., & Giles-Corti, B. (2017). The walkability planning support system: an evidence-based tool to design healthy communities (pp. 153-165). Springer International Publishing.

Chan, K., & Farber, S. (2020). Factors underlying the connections between active transportation and public transit at commuter rail in the Greater Toronto and Hamilton Area. *Transportation*, *47*(5), 2157-2178. https://doi.org/10.1007/s11116-019-10006-w

D'Orso, G., & Migliore, M. (2020). A GIS-based method for evaluating the walkability of a pedestrian environment and prioritised investments. *Journal of transport geography*, 82, 102555. https://doi.org/10.1016/j.jtrangeo.2019.102555

Gaglione, F., Cottrill, C., & Gargiulo, C. (2021). Urban services, pedestrian networks, and behaviors to measure elderly accessibility. *Transportation research part D: transport and environment*, *90*, 102687. https://doi.org/10.1016/j.trd.2020.102687

Koo, B. W., Guhathakurta, S., Botchwey, N., & Hipp, A. (2023). Can good microscale pedestrian streetscapes enhance the benefits of macroscale accessible urban form? An automated audit approach using Google street view images. *Landscape and Urban Planning*, *237*, 104816. https://doi.org/10.1016/j.landurbplan.2023.104816

Lee, J., Abdel-Aty, M., & Shah, I. (2019). Evaluation of surrogate measures for pedestrian trips at intersections and crash modeling. *Accident Analysis & Prevention*, *130*, 91-98. https://doi.org/10.1016/j.aap.2018.05.015

Li, Y., Yabuki, N., & Fukuda, T. (2023). Integrating GIS, deep learning, and environmental sensors for multicriteria evaluation of urban street walkability. Landscape and Urban Planning, 230, 104603. https://doi.org/10.1016/j.landurbplan.2022.104603

Li, Y., Yabuki, N., & Fukuda, T. (2023). Integrating GIS, deep learning, and environmental sensors for multicriteria evaluation of urban street walkability. *Landscape and Urban Planning*, *230*, 104603.https://doi.org/10.1016/j.landurbplan.2022.104603

Liu, Y., Yang, D., Timmermans, H. J., & de Vries, B. (2020). Analysis of the impact of street-scale built environment design near metro stations on pedestrian and cyclist road segment choice: A stated choice experiment. *Journal of transport geography,* 82, 102570. https://doi.org/10.1016/j.jtrangeo.2019.10250

Loo, B. P., & Lam, W. W. Y. (2012). Geographic accessibility around health care facilities for elderly residents in Hong Kong: A microscale walkability assessment. Environment and Planning B: Planning and Design, 39(4), 629-646. https://doi.org/10.1068/b36146

Loo, B. P., & Lam, W. W. Y. (2012). Geographic accessibility around health care facilities for elderly residents in Hong Kong: A microscale walkability assessment. *Environment and Planning B: Planning and Design*, *39*(4), 629-646. https://doi.org/10.1068/b36146

Moon, M., Lee, Y., Yu, K., & Kim, J. (2016). Optimized path finding algorithm for walking convenience of the people with reduced mobility. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and*

Cartography, *34*(3), 273-282. https://doi.org/10.7848/ksgpc.2016.34.3.273

Murgante, B., Borruso, G., Lapucci, A. (2009). Geocomputation and Urban Planning. In: Murgante, B., Borruso, G., Lapucci, A. (eds) Geocomputation and Urban Planning. Studies in Computational Intelligence, vol 176. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-89930-3_1

Ruiz-Padillo, A., Pasqual, F. M., Uriarte, A. M. L., & Cybis, H. B. B. (2018). Application of multi-criteria decision analysis methods for assessing walkability: A case study in Porto Alegre, Brazil. *Transportation research Part D: transport and environment*, 63, 855-871. https://doi.org/10.1016/j.trd.2018.07.016

S. Alfosool, A. M., Chen, Y., & Fuller, D. (2022). ALF–Score—A novel approach to build a predictive network–based walkability scoring system. Plos one, 17(6), e0270098. https://doi.org/10.1371/journal.pone.0270098

Southworth, M. (2005). Designing the walkable city. *Journal of urban planning and development,* 131(4), 246-257. https://doi.org/10.1061/(ASCE)0733-9488(2005)131:4(246)

Stipancic, J., Miranda-Moreno, L., Strauss, J., & Labbe, A. (2020). Pedestrian safety at signalized intersections: Modelling spatial effects of exposure, geometry and signalization on a large urban network. *Accident Analysis & Prevention*, *134*, 105265. https://doi.org/10.1016/j.aap.2019.105265

Su, S., Zhou, H., Xu, M., Ru, H., Wang, W., & Weng, M. (2019). Auditing street walkability and associated social inequalities for planning implications. Journal of transport geography, 74, 62-76. https://doi.org/10.1016/j.jtrangeo.2018.11.003

Suminski, R. R., Kelly, K., & Plautz, E. (2023). The associations between place-based measures of walkability and physical activity across a range of diverse streetscapes. Preventive Medicine, 169, 107454. https://doi.org/10.1016/j.ypmed.2023.107454

Valls, F., & Clua, Á. (2023). Modeling Barcelona sidewalks: A high resolution urban scale assessment of the geometric attributes of the walkable network. *PloS one*, *18*(7), e0284630. https://doi.org/10.1371/journal.pone.0284630

Yin, L. (2017). Street level urban design qualities for walkability: Combining 2D and 3D GIS measures. *Computers, Environment and Urban Systems, 64,* 288-296. https://doi.org/10.1016/j.compenvurbsys.2017.04.001