

Crowdsourcing and Sidewalk Data: A Preliminary Study on the Trustworthiness of OpenStreetMap Data in the US

Kazi Shahrukh Omar¹, Gustavo Moreira¹, Daniel Hodczak¹,
Maryam Hosseini², Fabio Miranda¹

¹ University of Illinois Chicago, ² New York University

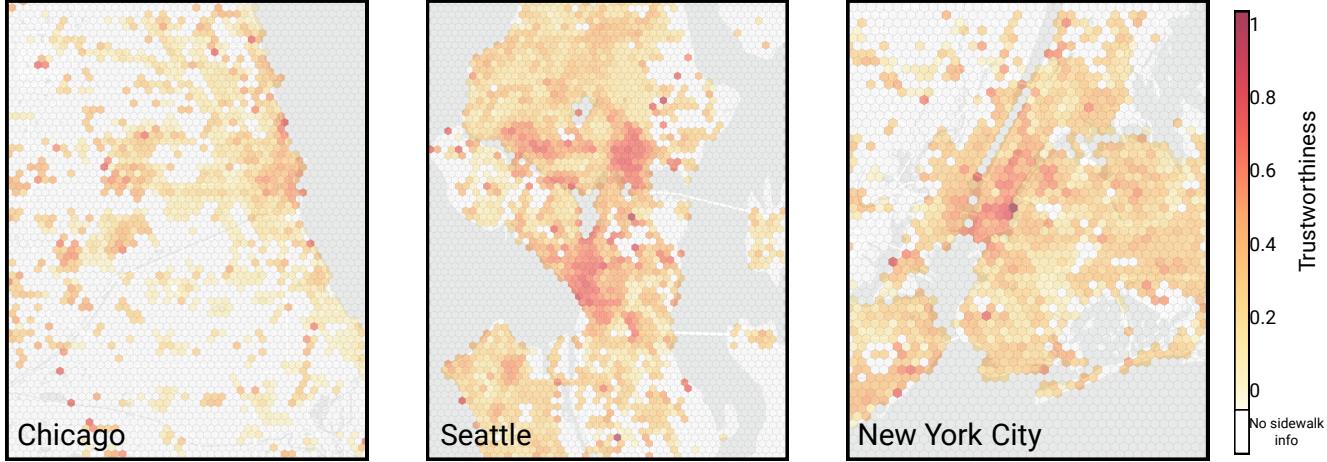


Figure 1. Spatial distribution of the average trustworthiness of *roads with sidewalk information and sidewalk geometries* in Chicago, Seattle and New York City.

Abstract

Sidewalks play a pivotal role in urban mobility of everyday life. Ideally, sidewalks provide a safe walkway for pedestrians, link public transportation facilities, and equip people with routing and navigation services. However, there is a scarcity of open sidewalk data, which not only impacts the accessibility and walkability of cities but also limits policymakers in generating insightful measures to improve the current state of pedestrian facilities. As one of the most famous crowdsourced data repositories, OpenStreetMap (OSM) could aid the lack of open sidewalk data to a large extent. However, completeness and quality of OSM data have long been a major issue. In this paper, we offer a preliminary study on the availability and trustworthiness of OSM sidewalk data. First, we compare OSM sidewalk data coverage in over 50 major cities in the United States. Then, we select three major cities (Seattle, Chicago, and New York City) to further analyze the completeness of sidewalk data and its features, and to compute a trustworthiness index leveraging historical OSM sidewalk data.

1. Introduction

Sidewalks are arguably the most important pedestrian-dedicated public spaces. They are the focal point of cities at the human scale, where the most widely used and environmentally sustainable form of transportation (walk-

ing/rolling) takes place [1, 2]. Sidewalks can significantly impact the everyday life of people, specifically for those relying on such infrastructure as the primary means of accessing public spaces [3–7] and public transit [8]. Well-designed sidewalks can also support local businesses and promote economic activities [9, 10]. Despite their prominent role in various pedestrian-level analysis, the available public data on sidewalks is significantly scarce and limited in time and geographical coverage [11]. Ironically, data collection and monitoring of motorized travel patterns and infrastructure have been practiced since the 1950s [12], and the collected data has been extensively used in transportation planning research, while collecting pedestrian-level data has only recently received some attention [13]. Furthermore, the existing datasets are collected mainly by resourceful cities, with substantial variation in the extent of data and attributes, and inconsistent methods from place to place. Such problems create significant barriers to conducting comparative studies, or data integration, across administrative borders [13–15]. Majority of the sidewalk inventories still lack important feature such as width, surface materials, curb ramps, tactile surface indicators, audio signals, etc., which are essential for specific mobility needs [15]. This lack of sidewalk data not only affects accessible routing in cities but also impacts how urban planners and policymakers could scrutinize mobility issues and the condi-

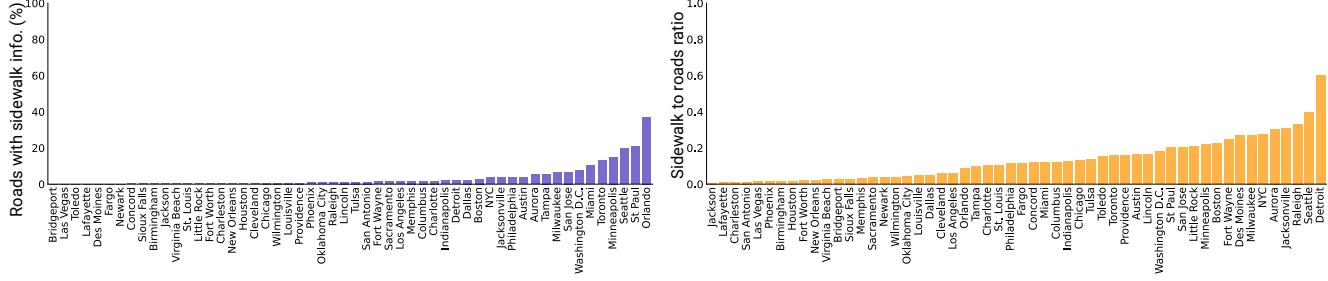


Figure 2. Comparing the coverage of sidewalk data in OSM for 54 cities in the US. Left: percentage of roads with sidewalk information. Right: ratio between sidewalks and roads. The majority of cities have less than 20% of roads with sidewalk information. In addition, only Seattle and Detroit have a sidewalk to road ratio above 0.4.

tion of sidewalks to make informed decisions [16]. Ongoing projects [4, 17–22] propose to standardize and improve sidewalk data coverage. Even so, researchers primarily rely on crowdsourced OpenStreetMap (OSM) data, the largest geospatial open-data initiative, with data covering not only streets and roads, but also buildings, points of interest, and other geographic entities [23].

Considering OSM’s importance in sidewalk studies, our primary goal is to assess the coverage and trustworthiness of OSM sidewalk data. While previous studies [24, 25] have assessed completeness of sidewalks, we propose to also assess the trustworthiness of the data by analyzing its history and provenance. Concretely, we compare the sidewalk coverage of OSM in 54 major cities in the United States and further expand our analysis in three major cities (Seattle, Chicago and New York City) with the spatial trustworthiness of the data. We then highlight possible research directions to mitigate some of the identified problems.

2. Related Work

OSM data has been used in various studies spanning different fields, including routing [26, 27], location-based services [28–30], traffic and transportation [31–34], energy modeling [35, 36], population estimation [37, 38], 3D city modeling [39–41], land cover use [42, 43], and emergency response management [44, 45]. Ongoing initiatives to improve sidewalk data, including Accessmap [18] and Open-Sidewalks [17], also utilize data from OSM. However, quality of OSM data has always been a major concern for both research and industrial purposes [46–54]. To tackle this problem, studies have been conducted to evaluate the quality and completeness of OSM data, focusing on different entities, such as roads [55–57], buildings [49, 50], and points of interest [58]. Properly assessing the quality of OSM data is a challenge, given that traditional approaches rely on the availability of official data (though even such data might have problems, including slow update rate). Alternative approaches then analyze the *evolution* of the data itself, assessing how the number of users editing, confirmations from different users, number of versions and rollbacks con-

tribute to the quality of OSM data [59–66]. Considering this, we propose to evaluate the use of trustworthiness as an index for OSM sidewalk data.

3. Study Area and OpenStreetMap Data

Between 2001 and 2019, the US built-up land - buildings, roads, and other infrastructures - increased by more than 14,000 square miles of new developments [67]. Today, more than 80% of the US population (around 309 million people) lives in urban areas [68]. Subsequently, there has been a growth in the availability of OSM data across the US, capturing its complex and diverse built environment. Our study first analyzes the availability of OSM sidewalk

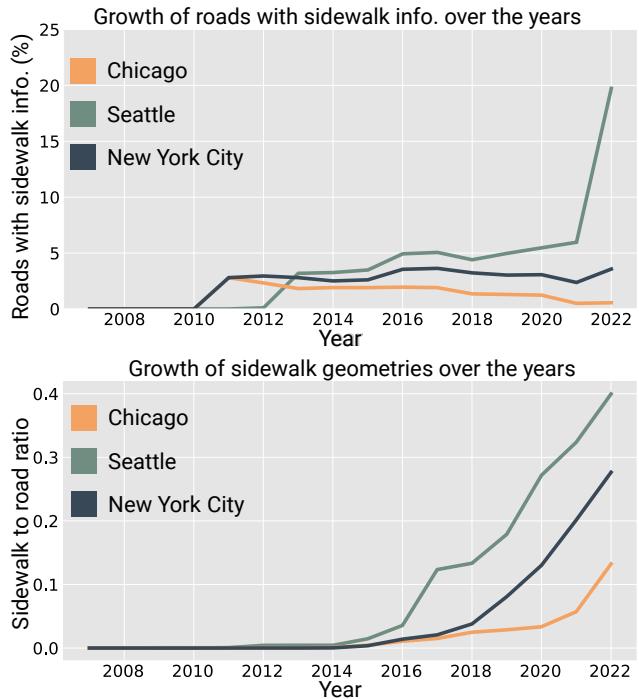


Figure 3. Growth of OSM sidewalk data in Chicago, Seattle, and NYC over the years. Top: growth of roads with sidewalk information. Bottom: growth of sidewalk / road ratio.

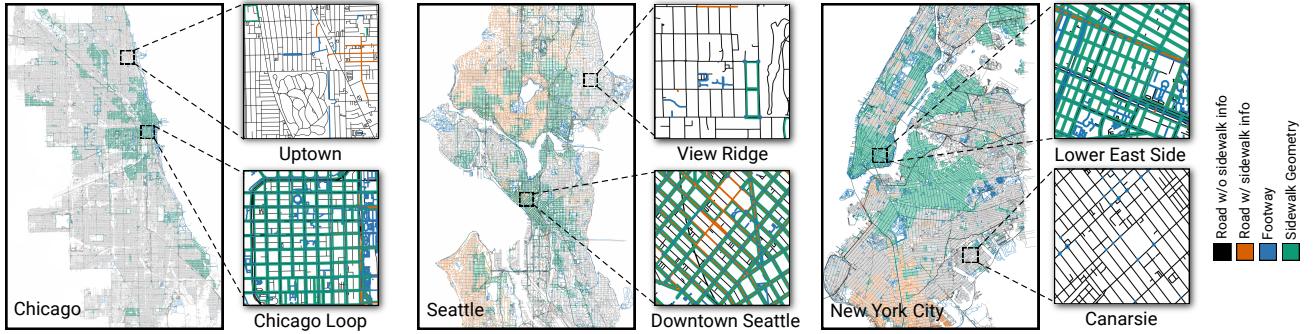


Figure 4. Coverage of *sidewalks* and *roads with sidewalk info* in Chicago, Seattle, and New York City (from left to right). Seattle OSM has comparably higher percentage of sidewalk geometries and roads with sidewalk information mapped than the other two cities. Also, Neighborhoods like *Chicago Loop*, *Downtown Seattle*, and Manhattan’s *Lower East Side* have good sidewalk coverage. On the other hand *Uptown*, *View Ridge*, and *Canarsie* neighborhoods only have a handful of mapped sidewalk geometries.

data (Sec. 4), followed by an analysis on the trustworthiness of the data (Sec. 5). First, we begin with an overview of the structure of the OSM data and how we use it.

The OSM data model is built upon three basic elements: *nodes* (points in space), *ways* (linear features and boundaries), and *relations* (multi-purpose entities that define relationship between two or more elements – nodes, ways, and/or other relations). OSM elements have *tags* attached to them that describe different physical features on the ground (e.g., roads, buildings). For example, *ways* that have the *highway* tag are used to define highways, roads, streets, or paths. In this paper, highway and road are used interchangeably. In our analyses, we focus on sidewalks and roads with sidewalk info. In OSM, these features can appear as:

1. Sidewalk as a refinement to a highway: The sidewalk is a property of an existing road. Elements with `highway=*` + `sidewalk=[both,right,left,yes]`. Here, we excluded OSM road types that cannot have sidewalks along them (i.e., `footway`, `escape`, `raceway`, `busway`, `bridleway`, `path`, `cycleway`, `construction`, `corridor`).
2. Sidewalk as a separate way: The sidewalk is described as its own separate geometry. Elements with `highway=footway` + `footway=sidewalk`.

In this paper, we make a distinction between the two types, given that sidewalk geometries offer a more spatially accurate representation of the pedestrian environment [69]. Figure 2 presents the coverage of sidewalk data in OSM for 54 cities, considering both sidewalks as a refinement to a highway and sidewalks as separate ways (i.e., distinct geometry). Even though there has been a growth in the availability of OSM sidewalk data over the years (Figure 3), there is still a lack of data regarding this pedestrian infrastructure.

4. Coverage of OSM Sidewalk Data

In our work, we measure the object and attribute coverage of OSM sidewalk data. For the object-level analysis, we evaluated the spatial coverage of the following sidewalk categories: roads without sidewalk information, roads with sidewalk information, footway and sidewalk geometry. Figure 4 shows the coverage of each category in Chicago, Seattle and New York City. We can notice how Seattle’s data surpasses the other two cities in terms of roads with sidewalk info (19.56%), and sidewalk to road ratio (0.4). Table 1 illustrates the statistics of OSM sidewalk data for all three cities, with Chicago and NYC both having a very low percentage of roads with sidewalk information (respectively 0.55% and 3.50%). Also noteworthy is that there are several regions in these cities with almost no sidewalk data (e.g., *Uptown*, *View Ridge*, and *Canarsie*).

For the attribute-level analysis, we assessed the number of additional attributes available in OSM sidewalk geometries (e.g., surface material, width, wheelchair accessibility). These features are of great importance in routing and navigation for people with mobility or vision impairments. Figure 5 shows the availability of different features in the chosen cities. The most commonly present feature in all three cities (surface material) varied between 40.5% in Seattle, 27% in NYC, and less than 10% in Chicago. Some of the attributes (such as kerb, smoothness, and width) were almost non-existent in all three cities.

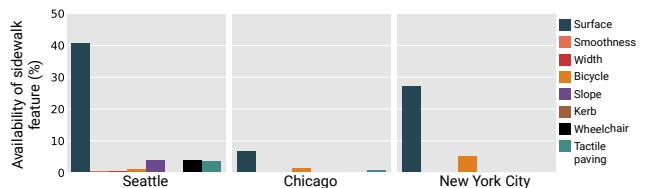


Figure 5. Availability of OSM sidewalk attributes. *Surface material* is the most commonly available feature in all three cities whereas *kerb*, *smoothness* and *width* are almost non-existent.

Table 1. Coverage of OSM sidewalk data for selected cities.

City	# of roads	# of sidewalk geometries	# of roads with sidewalk info	Percentage of roads with sidewalk info (%)	Sidewalk to road ratio
Chicago	134,598	17,492	746	0.55	0.12
Seattle	44,013	17,487	8,610	19.56	0.40
New York City	180,255	49,736	6,311	3.50	0.28

5. Trustworthiness of OSM Sidewalk Data

An important challenge in assessing the quality of OSM data is the lack of official data. Previous studies have then proposed to analyze the evolution of the data itself (history and provenance) [63, 65, 66], evaluating edit history [59], number of editors [60], number of rollbacks and direct confirmations from different users [61]. In our work, we make use of Alghanim et al.’s approach [66] to compute a trustworthiness index of OSM sidewalk data, based on direct, indirect and temporal components. Following their approach, trustworthiness is defined as follows:

$$T(f_i) = w_d * T_{dir}(f_i) + w_i * T_{ind}(f_i) + w_{time} * T_{time}(f_i)$$

where $T_{dir}(f_i)$, $T_{ind}(f_i)$ and $T_{time}(f_i)$ are the direct, indirect and temporal indicators of the $i - th$ version of feature f , and w_{dir} , w_{ind} , and w_{time} are weights balancing the components (0.5, 0.25, 0.25, respectively). A direct indicator depends on the feature version historical information. An indirect indicator models contributions that do not directly depend on the current feature version, but its neighboring spatial features. A temporal indicator models the impact of time on a feature trustworthiness, i.e., the longer a feature persists, the higher the chances it actually matches the real feature. For a more detailed discussion, we refer the reader to Alghanim et al. [66]. Figure 6 shows the trustworthiness index distribution for both *roads with sidewalk information* and *sidewalk geometries*, and Tables 2 and 3 highlight the percentage with *high* trustworthiness. Given these preliminary results, not only sidewalk coverage might be an issue, but also the trustworthiness of this data.

Table 2. Trustworthiness index of *roads with sidewalk info*.

City	T-index < 0.5 (%)	T-index ≥ 0.5 (%)
Chicago	75.64	24.36
Seattle	72.31	27.69
NYC	75.27	24.73

Table 3. Trustworthiness index of *sidewalk geometries*.

City	T-index < 0.5 (%)	T-index ≥ 0.5 (%)
Chicago	90.23	9.77
Seattle	78.09	21.91
NYC	83.0	17.0

6. Discussion and Future Work

In this study, we take steps towards understanding the quality OSM sidewalk data and evaluating both its cover-

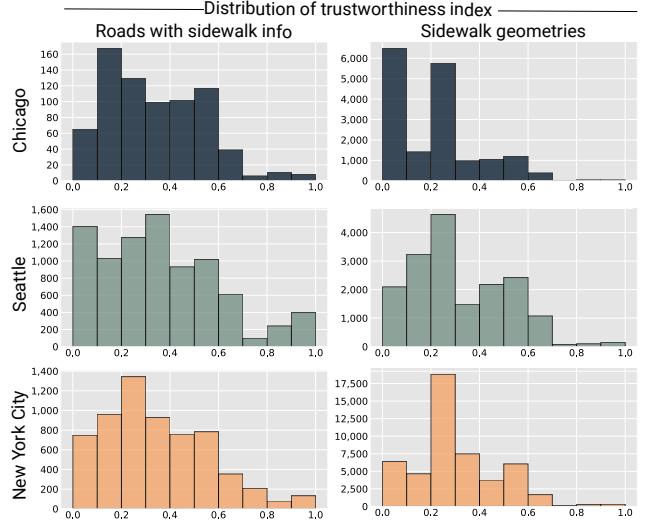


Figure 6. Distribution of trustworthiness index of *roads with sidewalk information* and *sidewalk geometries* in the selected cities.

age and trustworthiness, going beyond previous studies that only focused on coverage. The significance of sidewalks to everyday life of urban dwellers coupled with the lack of publicly available datasets describing their locations and attributes, underscore the importance of standardized datasets such as OSM, specifically in the absence of authoritative data. Although preliminary, our goal is to shed light on the availability of OSM sidewalk data, as well providing a method to analyze whether and to what extent the available OSM sidewalk data can be trusted. For example, our results show that even among the major and often resourceful cities in the US, there is a significant lack of OSM sidewalk data (Figure 2). Among the 54 cities analyzed, only 11% have more than 10% of their roads tagged with sidewalk information, and 80% of cities fall below 5%. In Orlando, ranked first, only 37% of the roads have any sidewalk-related tags.

Regarding attributes in the three selected cities, surface material has the highest availability, with 40.4% of Seattle’s sidewalks, 26.9% of NYC’s sidewalks, and 6.6% of Chicago’s sidewalks having this information. The remaining attributes were found in insignificant quantities or were absent in the select cities. This analysis further illustrates the inadequacy of the existing data in supporting comprehensive analysis of sidewalks [70]. In future work, we plan to further evaluate the measures adopted in this study, using authoritative data to validate the trustworthiness and correctness of the data. We will also investigate new techniques to extract important sidewalk attributes.

References

- [1] David S Vale, Miguel Saraiva, and Mauro Pereira. Active accessibility: A review of operational measures of walking and cycling accessibility. *Journal of transport and land use*, 9(1):209–235, 2016. 1
- [2] Erica Twardzik, Kate Duchowny, Amby Gallagher, Neil Alexander, Debra Strasburg, Natalie Colabianchi, and Philippa Clarke. What features of the built environment matter most for mobility? using wearable sensors to capture real-time outdoor environment demand on gait performance. *Gait & posture*, 68:437–442, 2019. 1
- [3] Han Qin, Kevin M Curtin, and Matthew T Rice. Pedestrian network repair with spatial optimization models and geocrowdsourced data. *GeoJournal*, 83(2):347–364, 2018. 1
- [4] Manaswi Saha, Michael Saugstad, Hanuma Teja Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, et al. Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019. 1, 2
- [5] Marcos AG Ferreira and Suely da Penha Sanches. Proposal of a sidewalk accessibility index. *Journal of Urban and Environmental Engineering*, 1(1):1–9, 2007. 1
- [6] Philippa Clarke, Jennifer A Ailshire, Michael Bader, Jeffrey D Morenoff, and James S House. Mobility disability and the urban built environment. *American journal of epidemiology*, 168(5):506–513, 2008. 1
- [7] Dori E Rosenberg, Deborah L Huang, Shannon D Simonovich, and Basia Belza. Outdoor built environment barriers and facilitators to activity among midlife and older adults with mobility disabilities. *The Gerontologist*, 53(2):268–279, 2013. 1
- [8] Mintesnot Woldeamanuel and Andrew Kent. Measuring walk access to transit in terms of sidewalk availability, quality, and connectivity. *Journal of Urban Planning and Development*, 142(2):040150191–0401501913, 2016. 1
- [9] Annette M Kim. The mixed-use sidewalk: Vending and property rights in public space. *Journal of the American Planning Association*, 78(3):225–238, 2012. 1
- [10] Yvonne Yen Liu, Patrick Burns, and Daniel Flaming. Sidewalk stimulus: Economic and geographic impacts of los angeles street vendors. *Economic Roundtable Report*, pages 1–29, 2015. 1
- [11] Shiloh Deitz, Amy Lobben, and Arielle Alferez. Squeaky wheels: Missing data, disability, and power in the smart city. *Big Data & Society*, 8(2):1–16, 2021. 1
- [12] Krista Nordback, Sirisha Kothuri, Theodore Petritsch, Peyton McLeod, Eliot Rose, and Hannah Twaddell. Exploring pedestrian counting procedures. a review and compilation of existing procedures, good practices, and recommendations. Technical report, FHWA-HPP-16-026, 2016. 1
- [13] Hugh Louch, Kim Voros, and Erin David. *Availability and Use of Pedestrian Infrastructure Data to Support Active Transportation Planning*. Number Project 20-05, Topic 50-10. 2020. 1
- [14] Shiloh Deitz. *Free Movement: Enhancing Open Data to Facilitate Independent Travel for Persons with Disabilities*. PhD thesis, University of Oregon, 2021. 1
- [15] Qing Hou and Chengbo Ai. A network-level sidewalk inventory method using mobile lidar and deep learning. *Transportation Research Part C: Emerging Technologies*, 119:102772, 2020. 1
- [16] Gary KW Chin, Kimberly P Van Niel, Billie Giles-Corti, and Mathew Knuiman. Accessibility and connectivity in physical activity studies: The impact of missing pedestrian data. *Preventive medicine*, 46(1):41–45, 2008. 2
- [17] Taskar Center for Accessible Technology (TCAT) at The University of Washington. Opensidewalks. <https://www.opensidewalks.com/>, 2021. Accessed: August 05, 2022. 2
- [18] Nicholas Bolten and Anat Caspi. Accessmap website demonstration: Individualized, accessible pedestrian trip planning at scale. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, pages 676–678, 2019. 2
- [19] Maryam Hosseini, Fabio Miranda, Jianzhe Lin, and Claudio T Silva. Citysurfaces: City-scale semantic segmentation of sidewalk materials. *Sustainable Cities and Society*, 79:103630, 2022. 2
- [20] Amin Mobasher, Haosheng Huang, Lívia Castro Degrossi, and Alexander Zipf. Enrichment of OpenStreetMap data completeness with sidewalk geometries using data mining techniques. *Sensors*, 18(2):509, 2018. 2
- [21] Tatiane Ferreira Olivatto, Fábio Noel Stanganini, and Edson Augusto Melanda. Drone imagery for OpenStreetMap sidewalk data enrichment. In *LBS 2019; Adjunct Proceedings of the 15th International Conference on Location-Based Services/Gartner, Georg; Huang, Haosheng*, pages 243–249. Wien, 2019. 2
- [22] Kauê Vestena, Silvana Camboim, and Daniel Santos. OSM Sidewalkreator-a QGIS plugin for automated sidewalk drawing for OSM. *Editors*, page 34, 2022. 2
- [23] Mordechai Haklay and Patrick Weber. OpenStreetMap: User-generated street maps. *IEEE Pervasive computing*, 7(4):12–18, 2008. 2
- [24] Amin Mobasher, Mohamed Bakillah, Adam Rousell, Stefan Hahmann, and Alexander Zipf. On the completeness of sidewalk information in OpenStreetMap, a case study of germany. In *The 18th AGILE International Conference on Geographic Information Science*, 2015. 2
- [25] Amin Mobasher, Yeran Sun, Lukas Loos, and Ahmed Loai Ali. Are crowdsourced datasets suitable for specialized routing services? case study of OpenStreetMap for routing of people with limited mobility. *Sustainability*, 9(6):997, 2017. 2

- [26] Dennis Luxen and Christian Vetter. Real-time routing with OpenStreetMap data. In *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems*, pages 513–516, 2011. 2
- [27] Soraia Felício, Joana Hora, Marta Campos Ferreira, Diogo Abrantes, Paulo Dias Costa, Camila Dangelo, Jorge Silva, and Teresa Galvão. Handling OpenStreetMap georeferenced data for route planning. *Transportation Research Procedia*, 62:189–196, 2022. 2
- [28] Blazej Ciepluch, Peter Mooney, Ricky Jacob, and Adam C Winstanley. Using OpenStreetMap to deliver location-based environmental information in ireland. *SIGSPATIAL Special*, 1(3):17–22, 2009. 2
- [29] Rajib Chandra Das and Tauhidul Alam. Location based emergency medical assistance system using OpenStreetMap. In *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, pages 1–5. IEEE, 2014. 2
- [30] Rajib Chandra Das, Parijat Prashun Purohit, Tauhidul Alam, and Mahfuzulhoq Chowdhury. Location based atm locator system using OpenStreetMap. In *The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014)*, pages 1–6. IEEE, 2014. 2
- [31] Michael Zilske, Andreas Neumann, and Kai Nagel. *OpenStreetMap for traffic simulation*. Technische Universität Berlin Berlin, Germany, 2015. 2
- [32] David Rieck, Björn Schünemann, and Ilja Radusch. Advanced traffic light information in OpenStreetMap for traffic simulations. In *Modeling Mobility with Open Data*, pages 25–34. Springer, 2015. 2
- [33] Chelcie Narboneta and Kardi Teknomo. OpenTripPlanner, OpenStreetMap, general transit feed specification: Tools for disaster relief and recovery. *arXiv preprint arXiv:1609.01472*, 2016. 2
- [34] Frank F Xu, Bill Y Lin, Qi Lu, Yifei Huang, and Kenny Q Zhu. Cross-region traffic prediction for China on OpenStreetMap. In *Proceedings of the 9th ACM SIGSPATIAL International Workshop on Computational Transportation Science*, pages 37–42, 2016. 2
- [35] Alaa Alhamwi, Wided Medjroubi, Thomas Vogt, and Carsten Agert. OpenStreetMap data in modelling the urban energy infrastructure: a first assessment and analysis. *Energy Procedia*, 142:1968–1976, 2017. 2
- [36] Jan Schiefelbein, Jana Rudnick, Anna Scholl, Peter Remmen, Marcus Fuchs, and Dirk Müller. Automated urban energy system modeling and thermal building simulation based on OpenStreetMap data sets. *Building and environment*, 149:630–639, 2019. 2
- [37] Mohamed Bakillah, Steve Liang, Amin Mobasher, Jamal Jokar Arsanjani, and Alexander Zipf. Fine-resolution population mapping using OpenStreetMap points-of-interest. *International Journal of Geographical Information Science*, 28(9):1940–1963, 2014. 2
- [38] Konštantín Rosina, Pavol Hurbánek, and Matej Cebečauer. Using OpenStreetMap to improve population grids in Europe. *Cartography and Geographic Information Science*, 44(2):139–151, 2017. 2
- [39] Martin Over, Arne Schilling, S Neubauer, and Alexander Zipf. Generating web-based 3D city models from OpenStreetMap: The current situation in Germany. *Computers, Environment and urban systems*, 34(6):496–507, 2010. 2
- [40] Harish Doraiswamy, Eleni Tzirita Zacharatou, Fabio Miranda, Marcos Lage, Anastasia Ailamaki, Cláudio T. Silva, and Juliana Freire. Interactive visual exploration of spatio-temporal urban data sets using urbane. In *Proceedings of the 2018 International Conference on Management of Data, SIGMOD ’18*, pages 1693–1696, New York, NY, USA, 2018. ACM. 2
- [41] Fabio Miranda, Harish Doraiswamy, Marcos Lage, Luc Wilson, Mondrian Hsieh, and Claudio T. Silva. Shadow accrual maps: Efficient accumulation of city-scale shadows over time. *IEEE Transactions on Visualization and Computer Graphics*, 25(3):1559–1574, March 2019. 2
- [42] Cidália Costa Fonte, Marco Minghini, Joaquim Patriarca, Vyron Antoniou, Linda See, and Andriani Skopeliti. Generating up-to-date and detailed land use and land cover maps using OpenStreetMap and GlobeLand30. *ISPRS International Journal of Geo-Information*, 6(4):1–22, 2017. 2
- [43] Cidália Costa Fonte, Joaquim António Patriarca, Marco Minghini, Vyron Antoniou, Linda See, and Maria Antonia Brovelli. Using OpenStreetMap to create land use and land cover maps: Development of an application. In *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications*, pages 1100–1123. IGI Global, 2019. 2
- [44] Melanie Eckle and João Porto de Albuquerque. Quality assessment of remote mapping in OpenStreetMap for disaster management purposes. In *Proceedings of the ISCRAM 2015 Conference*, pages 1–9. Citeseer, 2015. 2
- [45] Thiago Henrique Poiani, Roberto Dos Santos Rocha, Lívia Castro Degrossi, and Joao Porto De Albuquerque. Potential of collaborative mapping for disaster relief: A case study of OpenStreetMap in the Nepal earthquake 2015. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 188–197. IEEE, 2016. 2
- [46] Robert Hecht, Carola Kunze, and Stefan Hahmann. Measuring completeness of building footprints in OpenStreetMap over space and time. *ISPRS International Journal of Geo-Information*, 2(4):1066–1091, 2013. 2
- [47] Lindy-Anne Siebritz. Assessing the accuracy of OpenStreetMap data in South Africa for the purpose of integrating it with authoritative data. Master’s thesis, University of Cape Town, 2014. 2
- [48] SNRM Husen, NH Idris, and MHI Ishak. The quality of OpenStreetMap in Malaysia: A preliminary assessment. *International Conference on Geomatics and Geospatial Technology*, pages 291–298, 2018. 2
- [49] Maria Antonia Brovelli and Giorgio Zamboni. A new method for the assessment of spatial accuracy and completeness of OpenStreetMap building footprints. *ISPRS International Journal of Geo-Information*, 7(8):289, 2018. 2
- [50] Melih Basaraner. Geometric and semantic quality assessments of building features in OpenStreetMap for some areas

- of Istanbul. *Polish Cartographical Review*, 52(3):94–107, 2020. 2
- [51] Chen Wai Hoong. *Assessing the quality of OpenStreetMap building data in Singapore*. PhD thesis, National University of Singapore, 2020. 2
- [52] Renoy Girindran, Doreen S Boyd, Julian Rosser, Dhanya Vijayan, Gavin Long, and Darren Robinson. On the reliable generation of 3D city models from open data. *Urban Science*, 4(4):1–21, 2020. 2
- [53] Sylwia Borkowska and Krzysztof Pokonieczny. Analysis of OpenStreetMap data quality for selected counties in Poland in terms of sustainable development. *Sustainability*, 14(7):3728, 2022. 2
- [54] Hao Wu, Anqi Lin, Keith C Clarke, Wenzhong Shi, Abraham Cardenas-Tristan, and Zhenfa Tu. A comprehensive quality assessment framework for linear features from volunteered geographic information. *International Journal of Geographical Information Science*, 35(9):1826–1847, 2021. 2
- [55] Musfira Jilani, Padraig Corcoran, and Michela Bertolotto. Multi-granular street network representation towards quality assessment of OpenStreetMap data. In *Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Computational Transportation Science*, pages 19–24, 2013. 2
- [56] Musfira Jilani, Padraig Corcoran, and Michela Bertolotto. Automated highway tag assessment of OpenStreetMap road networks. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 449–452, 2014. 2
- [57] Masoud Minaei. Evolution, density and completeness of OpenStreetMap road networks in developing countries: the case of Iran. *Applied geography*, 119:102246, 2020. 2
- [58] Stefan Steiniger, Mohammad Ebrahim Poorazizi, Daniel R Scott, Cristian Fuentes, and Ricardo Crespo. Can we use OpenStreetMap POIs for the evaluation of urban accessibility? In *International Conference on GIScience Short Paper Proceedings*, volume 1, pages 1–5, 2016. 2
- [59] Peter Mooney and Padraig Corcoran. How social is OpenStreetMap. In *Proceedings of the 15th association of geographic information laboratories for Europe international conference on geographic information science, Avignon, France*, pages 24–27, 2012. 2, 4
- [60] Maurizio Napolitano and Peter Mooney. MVP OSM: A tool to identify areas of high quality contributor activity in OpenStreetMap. *The Bulletin of the Society of Cartographers*, 45(1):10–18, 2012. 2, 4
- [61] Carsten Keßler, Johannes Trame, and Tomi Kauppinen. Tracking editing processes in volunteered geographic information: The case of OpenStreetMap. In *Identifying objects, processes and events in spatio-temporally distributed data (IOPE), workshop at conference on spatial information theory*, volume 12, pages 6–8, 2011. 2, 4
- [62] Carsten Keßler and René Theodore Anton de Groot. Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. In *Geographic information science at the heart of Europe*, pages 21–37. Springer, 2013. 2
- [63] Fausto D’Antonio, Paolo Fogliaroni, and Tomi Kauppinen. VGI edit history reveals data trustworthiness and user reputation. *Proceedings of the AGILE’2014 International Conference on Geographic Information Science*, pages 1–5, 2014. 2, 4
- [64] Xiaoguang Zhou and Yijiang Zhao. A version-similarity based trust degree computation model for crowdsourcing geographic data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41:327–333, 2016. 2
- [65] Paolo Fogliaroni, Fausto D’Antonio, and Eliseo Clementini. Data trustworthiness and user reputation as indicators of VGI quality. *Geo-spatial Information Science*, 21(3):213–233, 2018. 2, 4
- [66] Amerah Alghanim, Musfira Jilani, Michela Bertolotto, and Gavin McArdle. Leveraging road characteristics and contributor behaviour for assessing road type quality in OSM. *ISPRS International Journal of Geo-Information*, 10(7):436, 2021. 2, 4
- [67] Zach Levitt and Jess Eng. Where america’s developed areas are growing: ‘way off into the horizon’. <https://www.washingtonpost.com/nation/interactive/2021/land-development-urban-growth-maps/>, 2021. Accessed: August 06, 2022. 2
- [68] Bendix Aria. A stunning new video shows how cities exploded onto the us map over the last 200 years. <https://www.businessinsider.com/us-urban-development-density-last-two-centuries-2020-6>, 2020. Accessed: August 06, 2022. 2
- [69] Maryam Hosseini, Iago B. Araujo, Hamed Yazdanpanah, Eric K. Tokuda, Fabio Miranda, Claudio T. Silva, and Roberto M. Cesar Jr. Sidewalk measurements from satellite images: Preliminary findings. In *2nd Spatial Data Science Symposium*, 2021. 3
- [70] Maryam Hosseini, Mikey Saugstad, Fabio Miranda, Andres Sevtsuk, Claudio T Silva, and Jon E Froehlich. Towards global-scale crowd+ ai techniques to map and assess sidewalks for people with disabilities. *arXiv preprint arXiv:2206.13677*, pages 1–6, 2022. 4