European Journal of Geography

The publication of the European Journal of Geography (EJG) is based on the European Association of Geographers' goal to make European Geography a worldwide reference and standard. Thus, the scope of the EJG is to publish original and innovative papers that will substantially improve, in a theoretical, conceptual, or empirical way the quality of research, learning, teaching, and applying geography, as well as in promoting the significance of geography as a discipline. Submissions are encouraged to have a European dimension. The European Journal of Geography is a peer-reviewed open access journal and is published quarterly.

Research Article

OSM Sidewalkreator: A QGIS plugin for an automated drawing of sidewalk networks for OpenStreetMap

Kauê de Moraes Vestena[1,✉], Silvana Philippi Camboim[2] & Daniel Rodrigues dos Santos[3],

[1] Federal University of Paraná: Curitiba, Paraná, Brazil
[2] Universidade Federal do Paraná: Curitiba, Paraná, Brazil
[3] Federal University of Paraná: Curitiba, Paraná, Brazil
✉ Correspondence: kauemv2@gmail.com

Received: 19/09/2023
Revised: 01/12/2023
Accepted: 09/12/2023
Published: 12/12/2023

DOI: 10.48088/ejg.k.ves.14.4.066.084
ISSN: 1792-1341

Abstract: Sidewalks are a relevant part of the living space in urban environments but are still rarely mapped. In recent years, the mapping of sidewalks has grown in importance among the OSM and academic communities as a matter of concern for many UN SDGs. To cover this gap, we propose a GitHub-hosted, fully open-source QGIS Plugin entitled "OSM SidewalKreator" to automatically draw the geometries of sidewalks for OSM crossings and curb crossing interfaces. The plugin workflow encompasses the steps of input area selection; data fetching, data cleaning, sidewalk geometries generation; crossings and kerbs generation; optional sidewalk splitting; and data exporting. Furthermore, the tool gives the user the capacity to have control over the process. Our tests revealed that the proposed method embodied by the plugin surpasses the manual process in many contexts, highlighting completeness, topological and thematic accuracies. We conclude that deepening, improving, and increasing the amount of open sidewalk mapping, mainly in the widely available OSM, can be a valuable asset to improve the development of accessibility and mobility worldwide.

Keywords: Pedestrian Networks; Geographic Information Systems; Open Source Software and Data; OpenStreetMap; Urban Mobility

Highlights:

- SIdewalks and other footpaths are generally rarely mapped despite their huge importance
- OSM SIdewalKreator QGIS PLugin is a product of the research, being a valuable support tool
- A deep analysis compared sidewalks produced with two methods, endorsing the method's strengths

1. Introduction

Considering urban scenarios, there are three types of pedestrian-centered paths (US Department of Transportation, 2013): walkways, paved shoulders, and sidewalks. For example, in New York City, due to its more than 10,000 kilometers of roads, there are more than 19,000 kilometers of sidewalks (NYC DOT, 2023). In most urban scenarios, sidewalks are the most prominent pedestrian facility (Kim, 2015; Rodriguez et al., 2015)

Sidewalks are a relevant part of the urban environment. The existence of sidewalks and their condition is fundamental to locomotion and is critically important in mobility groups such as cyclists, wheelchair users, blind people, elders, and children. Also, the displacement along sidewalks can ensure safety from traffic, contributing to the well-being of citizens (US Department of Transportation, 2013). Sidewalks are also valuable living spaces, being in some cases almost the only ones available, sharpening their commonplace where many actors (pedestrians, street vendors, urban vegetation, street furniture, among others) continuously contest their usage rights (Kim, 2015).

Simultaneously, promoting pedestrian mobility can contribute to the United Nations (UN) Sustainable Development Goals (SDGs) (UN, 2015). For instance, it can help reduce greenhouse gas emissions (Goal 13: Climate Action) and encourage a more active daily lifestyle (Goal 3: Good Health and Well-being). Leśniewski et al. (2021) take this idea further, citing a report that asserts sustainable transport may contribute to all 17 SDGs.

Despite their importance, sidewalk data is often missing from both authoritative data and modern mapping platforms, thus being deemed as a secondary feature (Rhoads et. al., 2023). This lack of data is a significant problem that directly impacts the scope of mobility and urban design studies (Froehlich et al., 2022), which are frequently disregarding important themes such as pedestrian safety (Bustos et al., 2021). There is also a reported lack of quality in sidewalk data. For example, in a study with data from 178 US cities, only 20% had sidewalk data, and only half published curb ramp locations or other details like condition, obstructions, materials, and others (Hosseini et al., 2023). In addition, Bartzokas-Tsiompras & Photis (2020) found a negative correlation between walkability and ethnic diversity in Berlin, meaning that often adequate urbanistic environments are not universally accessible yet.

However, sidewalks are often in poor condition and lack basic accessibility features, such as adequate width, the absence of barriers, and the presence of ramps, resulting in poor walkability, such as empirically shown in Bartzokas-Tsiompras et. al. (2021) and Bartzokas-Tsiompras et. al. (2023). This scenario requires significant efforts to improve accessibility. (Pereira et al., 2017; Fidalgo et al., 2021). Considering this context, the mapping of sidewalks becomes even more critical.

OSM is the most widely used user-contributed open cartographic data initiative (Mooney & Minghini, 2017; Neis & Zipf, 2012; Sehra et al., 2013) that can hold accurate data (Hakay, 2010). OpenStreetMap is a global geospatial data platform with a large amount of data on mobility.

Therefore, incrementing with the pedestrian dimension of such a dataset can be very beneficial, using the power of local knowledge to enrich the base with essential data for the accessibility and walkability of cities.

The current method of mapping pavement in OSM has limitations, as it relies solely on the road axis as the primary geometry. This approach may be overly simplistic and less helpful in understanding accessibility. A better strategy would involve creating separate geometries for sidewalks, offering more detailed information and insight into pedestrian infrastructure accessibility. This approach would also enable a comprehensive understanding of the pedestrian network's topology, its relationship to roads and intersections, and the inclusion of essential elements like ramps, surface conditions, and obstructions.

However, mapping geometries individually can be complex and time-consuming. To address this issue, this research proposes a tool to aid platform contributors in the QGIS software environment. A QGIS plugin called "OSM SidewalKreator" has been developed to automate specific procedures, guiding contributors in creating sidewalk network geometries (SNG). The workflow of this research is depicted in Figure 1, the relevant concepts are introduced at sections 4 and 5.

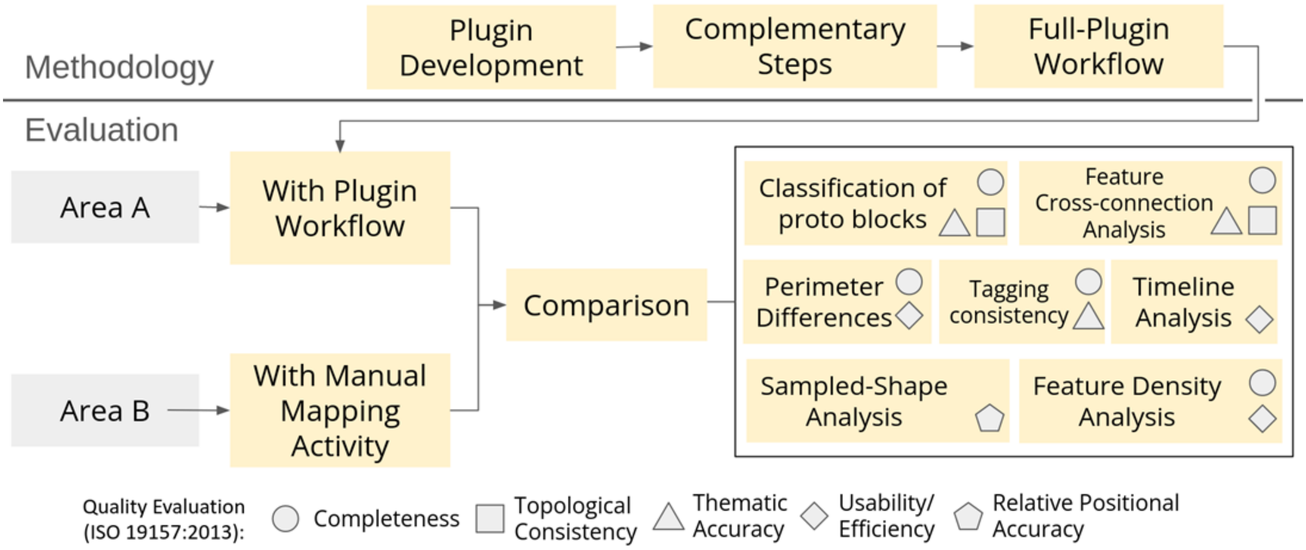

Figure 1. Employed Research Workflow

2. Related Work

Ertz et al. (2021) tested a framework based on a mobile app in order for volunteers to propose the best walkable paths throughout the city directly, manually. Jiang et al. (2021) used a semi-automated pipeline on QGIS to generate road crossings only. Gjeruldsen (2021) used machine learning to generate a Neural Network fed using OSM features to predict SNGs, with good results in areas from developed countries. Mobasheri et al. (2018) used data mining techniques to create SNG based on GNSS tracks given by wheelchair users. Kayama and Yairi (2007) used edge analysis on terrestrial images rectified as horizontal maps to detect sidewalk borders. Finally, Saha et al. (2015) created a platform for crowdsourcing information about sidewalks based on Google Street View imagery.

Most methods found in literature try to derive SNGs from data not connected to the streets but utterly connected to sidewalks. Most automated methods try to make some generalization to derive sidewalks. The description of sidewalk networks was also addressed in works such as Rhoads et. al. (2023).

Some OSM groups focus on mapping SNGs, such as the Open Sidewalks Initiative (University of Washington, 2016). This initiative serves as both a community and a project, offering dedicated mapping and a comprehensive scheme for pedestrian-centric mapping. However, as of 2023, all their methods still rely solely on the manual drawing of SNGs.

3. Sidewalk Mapping and OpenStreetMap

One criticism against the OSM data model concerns its loose mapping standards imbued in the principles of OSM guidelines: "The community agrees on certain key and value combinations for the most commonly used tags, which act as informal standards. However, users can create new tags to improve the style of the map or to support analyses that rely on previously unmapped attributes of the features" (OSM Wiki, 2023). Hartwig et al. (2015) and Ferster et al. (2019) present many ways bicycle paths can be mapped, exemplifying this topic. There are two major categories: separated ways (A) and properties (B) of the road (when adjacent). Similarly, the categories for sidewalks are the same, as presented by OSM Wiki (2016), Omar et al. (2022), and Biagi et al. (2020).

Figure 2 depicts these two available schemes. In schema A), we have three types of geometries to be represented: sidewalk paths at their actual axis, tagged as highway=footway + footway=sidewalk; crossings as highway=footway + footway=crossing; and access points as barrier=kerb. For B), one only puts if there is an allegedly parallel and straight sidewalk by adding the tag sidewalk=yes/no/left/right; one can add tagged nodes to aggregate the information capable of being represented at crossings.

There is still an open debate about which of the two ways best represents sidewalks in the OSM community. Some users claim in favour of the tagging scheme, arguing that over-representation can pollute the map and create unnecessary complexity (OSM Wiki, 2011). On the other

https://doi.org/10.48088/ejg.k.ves.14.4.066.084

hand, Biagi et al. (2020) and many OSM users nowadays (OSM Wiki, 2016) have been showing the representation of sidewalks as separated geometries as allegedly their best representation in OSM. There are many listed advantages (OSM Wiki, 2016): crossings may be represented as lines, with the kerb interfaces as points (8 in a regular 4-way intersection); the actual traversing length will be represented appropriately (as it will include block corners and crossings); independence from the digitizing direction, as "left" and "right" may swap if someone inverts the way direction; in tagging schema the actual sidewalk even when parallel, could be anywhere between road and first in-block feature (like a fence/wall or building, as depicted in Figure 2); ease of representation for pedestrian islands.

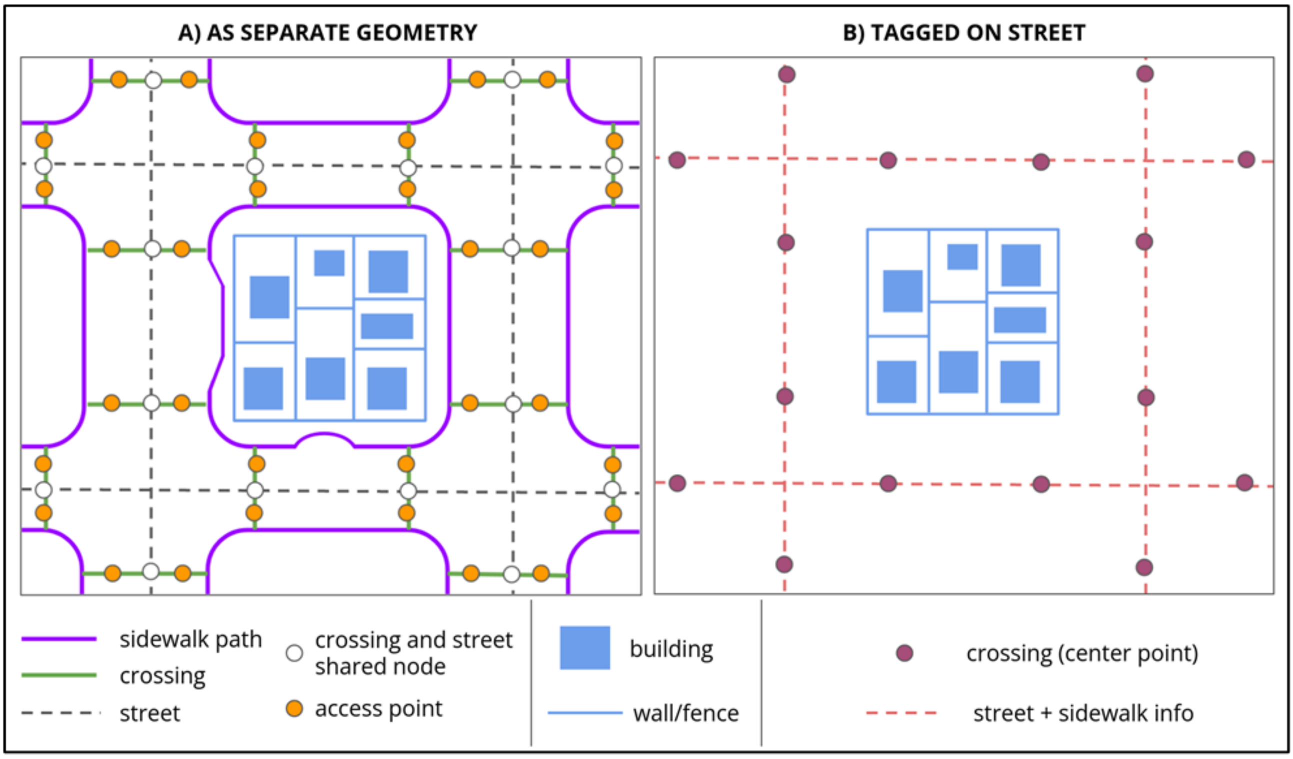

Figure 2. Possible Representations for Sidewalks in OSM

Moreover, some cases cannot be represented correctly using the tag scheme or will need some cumbersome solutions, as it shall represent the portion of the sidewalk orthogonally projected from the street, with properties inserted using compound tags such as "sidewalk:left/right:surface=*", with sidedness being of uttermost importance. Therefore, if a property differs on both sides, one may need to split the highway accordingly to represent it correctly, as shown in Figure 3. There are also other issues, e.g. geometric properties such as the distance from the sidewalk to the street will stay unclear. Another advantage is the correct topological representation of street furniture, considering, for example, if such a feature is on the sidewalk's left or right side, as shown in Figure 3.

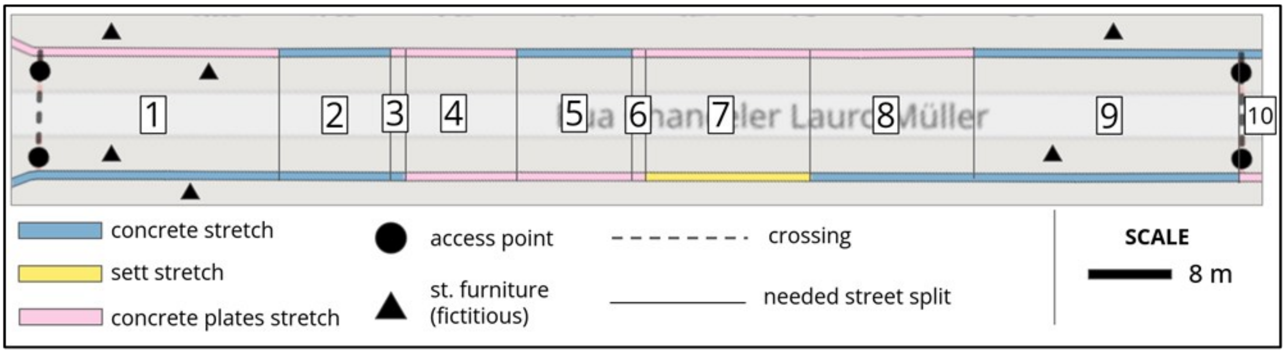

Figure 3. The necessity of multiple splits with sidewalks as just tags of the street

The representation of footways in OpenStreetMap (OSM) remains limited, regardless of the data representation method chosen. As of March 2023, our data extraction from Taginfo (OSM Contributors, 2023), a platform designed to provide statistical insights into OSM tags, revealed that the 'highway=*' tag, which designates transportation features, identifies approximately 215 million features in total. Within this dataset, a subset of 19.3 million (8.2% of the total) is tagged with 'highway=footway', indicating pedestrian paths. Further analysis of the 'footpath' tag key shows that only 6.2 million paths (2.88% of the total) are categorized as such, with 57.5% designated as 'sidewalk' and 41.3% as 'crossing'.

In the alternative scheme using the 'sidewalk=*' tag, we found only 3.2 million ways (1.6% of the total) classified as sidewalks. Given that a significant proportion of these features are in urban environments (Neis & Zielstra, 2014; Camboim et al., 2015), where roads typically have

https://doi.org/10.48088/ejg.k.ves.14.4.066.084

sidewalks (meaning that ideally, every kilometer of road should have almost two kilometers of sidewalks), it is clear that sidewalks are underrepresented in both data schemes.

Historically, this under-representation has been a persistent issue. Mobasheri et al. (2015) reported that in Berlin, at the time of their publication, only 5.6% of streets were tagged with 'sidewalk=*'. Even two years later, this figure had increased only slightly to 8.2% (Mobasheri et al., 2017).

Unfortunately, manually drawing sidewalks and crossings is time-consuming and can be error-prone. This effort can be inferred by examining the OpenSidewalk's Tasking Manager, an instance of a tool for managing collective mapping efforts by subdividing an area of interest into smaller square areas called "tasks" (Taskar Center, 2021). In its most near-completion project (Taskar Center, 2021), each task has taken an average of 9.4 minutes to be mapped but 22.7 minutes to be validated. Thus, considering just the mapping of crossings, for the 1046 existing tasks, it will take approximately 163.9 hours of mapping and 395.7 hours of validation. This total encompasses an area of just approximately 6.17 km2, only 0.65% of the urban area of São Paulo, for example.

4. Plugin Design and Operation

As previously mentioned, to help cover this gap in OSM sidewalk data, we propose a Github-hosted, fully open-source QGIS Plugin entitled "OSM SidewalKreator" (Vestena, 2021), which aims to automatically draw for OSM the geometries of sidewalk networks, along with the basic descriptive tags. The plugin, except when specifically stated otherwise by OSM data, works with the global presence of sidewalks and crossings as a basic premise, nevertheless, the methodology includes interaction windows to enable manual correction, if necessary.

The workflow of the plugin involves the use of external tools, as shown in Figure 4. The process begins by retrieving the base data from the OSM DB using the Overpass API. Then, the main stage of the process is performed in QGIS, where the user can interact with the data and trigger processing steps through the plugin's friendly GUI. The front end of the GUI is built on top of the QT library, and the back end is written in Python using the PyQGIS library. Finally, the output data generated by the plugin can be integrated with the existing data in the OSM DB through a conflation process that runs in JOSM, an advanced and extensible graphical OSM editor (OSM Germany, 2023), allowing the user to update the OSM DB.

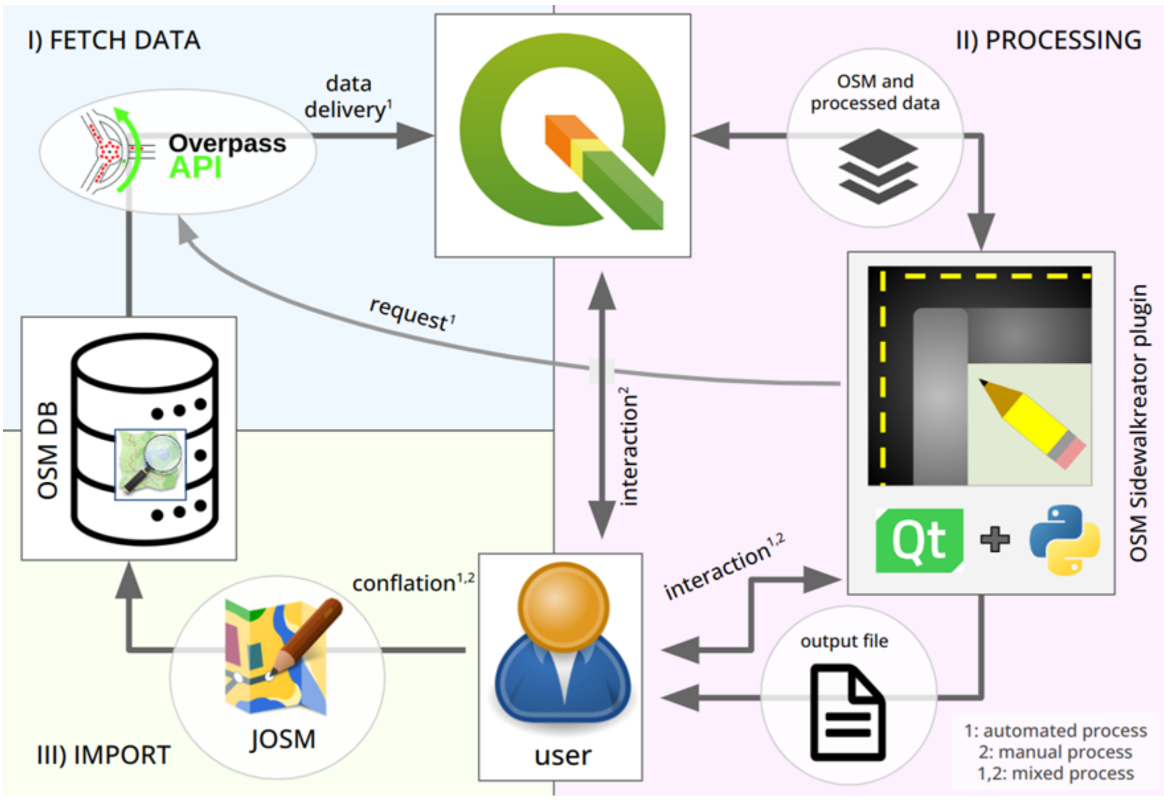

Figure 4. OSM SidewalKreator Full Workflow

The GUI-user interaction allows the user to control the partial outputs using QGIS GUI tools. The GUI includes action buttons that guide the user throughout the process. Once a button is pressed, it becomes deactivated, and the next button in the sequence is enabled for the user to trigger the next step. Figure 4 depicts the GUI, which contains six interaction elements highlighted in different colors. The green elements represent extra resources, such as base imagery and adding buttons. The red elements represent processing buttons, and the yellow elements represent the hint board. The hint board guides the user on the best moment to interact with the data beyond what is presented in the GUI.

Considering the user-plugin interaction, it happens through the plugin GUI, encompassing the seven following steps, most triggered by processing buttons:

1. Input area selection: The user can select a valid polygon feature, transforming its coordinates to WGS84 if needed.

2. Data fetching: As shown in Figure 4, the plugin calls the Overpass API to download the interest data from OSM DB. The primary interest data are streets, identified on OSM with the tag “highway=*”. Optionally, it can include data from both buildings and addresses. Each downloaded category is turned into a QGIS vector layer. In order to optimize the distortion issues on lengths, all data is converted to a local Transverse Mercator projection generated with its center-meridian passing through the input area polygon centroid. The user can freely edit OSM data starting at this step. However, this is recommended only after step 3.
3. Data Cleaning and Intersection Generation: All highway=* tag values available on fetched data are displayed in a table view. The user can then modify the default width values for each category. These values have a huge impact on the result, currently, they may be set empirically or statistically, with no built-in method to automatically compute them e.g. from aerial imagery, this holds for all parameters that must be set in the whole of the plugin. The values can be set according to the local reality. These values will be used for all features with that value without a proper width=* tag. All categories set below 0.5m are thrown out, keeping away highway=* tag values that aren’t streets (like footways or paths). The 'split at intersections' operation divides the remaining streets at road crossings. An optional iterative algorithm is provided for the purpose of excluding dead-end segments, entailing the exclusion of segments connected to only one other segment, repeating up to a user-set number of iterations.
4. Sidewalk Geometries Generation: At this stage, users have the option to customize the lengths of street segments before triggering the process. The area enclosed by streets is then converted into polygons and stored in a separate layer named "proto blocks." This step identifies complete blocks within the area of interest while discarding incomplete ones. The generation of sidewalks begins with buffering the street segments. The buffer width is calculated as half of the width determined in step 3, with an additional user-defined distance if the aim is to create sidewalk axes rather than curbs. If buildings are within this distance, the width is adjusted (before buffering) to be the distance to the nearest building minus the user-set "minimal distance to buildings." In cases where buildings are exceptionally close to the street, the width defaults to the "minimal width" set by the user. Following the buffering of street stretches, all generated rectangles are dissolved into a single shape. A sequence of one positive and one negative buffer is then applied, each with a defined corner curve radius, to round the corners. If a curve radius of zero is chosen, the corners will remain angular. The inner holes extracted from the dissolved polygon represent the sidewalk geometries, which are stored in a layer named "sidewalks." This layer includes four indicators in its attribute table: inner area, perimeter, area-to-perimeter ratio, and the square root of the area-to-perimeter ratio. These indicators assist users in identifying and excluding, if necessary, layer lines within "sidewalks" that do not represent actual sidewalks, such as traffic islands or roundabouts. To address the integration with existing sidewalk data, there are two implemented mechanisms:
   i. Inside proto blocks where more than 50% of the perimeter extension numerically overlaps with the perimeter of existing footways in the OSM database, the generated sidewalks are automatically excluded, thus assuming they have already been mapped; and
   ii. If a tag indicates the absence of sidewalks on one or both sides (either sidewalk=left/right/no or sidewalk:both/right/left=no), then a so-called “exclusion zone” is generated through a buffer process with an extra 10% of margin, ending up with overlap part of sidewalk being excluded. The geometries affected by this process may demand some extra review at the end of the whole process.
5. Crossings and kerb generation: An intersection can be created for each road segment's endpoint if the segment is connected to at least two other segments, given that ordinary road splits can occur at arbitrary points. A point is generated inward along the line if there is a crossing. The distance is calculated as half of the width of the intersecting road plus the curve radius and an additional fixed distance set by the user. However, this additional distance cannot exceed half the length of the road segment. A 1-meter buffer is generated around each newly created point, and a test is conducted against a dissolved version of the "proto blocks" layer. A road crossing is only generated if the small buffer circle is entirely contained within the larger polygon. This criterion ensures that sidewalks exist on both sides, nevertheless, this also means that proto blocks without inner sidewalks, such as the ones with eventual removals, the crossings wouldn't be generated. At the user's discretion, the crossing generation process can begin orthogonally or parallel to a transversal road, establishing a direction vector and its opposite. Both vectors are extended until an intersection is encountered with the "sidewalks" layer, where they are clipped. Subsequently, kerb access points are generated according to the user's specified percentage, both as points of crossings (resulting in a fixed count of 5 points per crossing) and in a distinct layer named "kerbs." It is worth noting that the generation process may lead to inconsistent or duplicated crossings. Like step 4, the user has the option to delete them with the assistance of several indicators, including crossing length, the difference between road width and actual length, a user-defined tolerance threshold (with an automatic exclusion option), and proximity to the nearest center point of another crossing, within a range of up to 50 meters.
6. Sidewalk splitting: In this optional step, the user can subdivide the sidewalks into smaller segments, each of which may have distinct attributes to accommodate various real-world conditions. Four options are available for this adjustment: the optimal, using the in-block Voronoi polygons concerning building centroids and addresses, by a maximum stretch length, by the number of segments, or only at corners. Users have the flexibility to select one of the last three options individually or in combination with the Voronoi-based adjustment, or they can choose not to split the blocks at all.
7. Data export: This process begins with data cleanup, which includes deleting all auxiliary fields created during the process. Fields corresponding to the respective OSM tags mentioned earlier are generated for each interest layer (sidewalks, crossings, and kerbs). Any loose kerb access points resulting from excluded crossings are removed. Subsequently, the interest layers are converted back to the WGS84 coordinate system and exported as a single .geojson file. This file can be directly opened in JOSM for further editing. Additionally, auxiliary files are generated to facilitate the inspection of internal processes, create a pre-tailored comment for the OSM changeset, and provide a .json dump containing the user-input parameters.

The next part of the methodology (a "step 8", external to the plugin) is the importation and conflation process, both happening at the JOSM editor. The first term refers to this whole part that begins by opening the output file and ends with uploading data on OSM. The second term, defined by Lei (2020) as "the process of matching counterpart features from two or more data sources in order to combine and better utilize information in the data", refers to the most critical part of this step, in this case checking possible "interactions" between the created features and already existing ones, like the mandatory creation of the intersection points between crossings and existent roads for proper representation of their topological relations, done automatically by JOSM tools. The user can also adjust created geometries to comply with satellite imagery for better positioning or to adapt generated geometries in case of irregular sidewalk shape, as in the example of a parking entrance shown in Figure 6. In addition, it may be necessary to create links between existing in-block footpaths and sidewalks, with other minor necessary edits on the data in agreement with pre-existing data, these are examples of manual procedures that still may need to be carried out by the user, although, as mentioned before, some conflation procedures can be undertaken by optimized tools at the editor. Finally, after the last step, the user can upload the new and eventually modified data to OSM DB, as mentioned above. The visual depiction of the plugin procedures is available as a workflow presented at Figure 7.

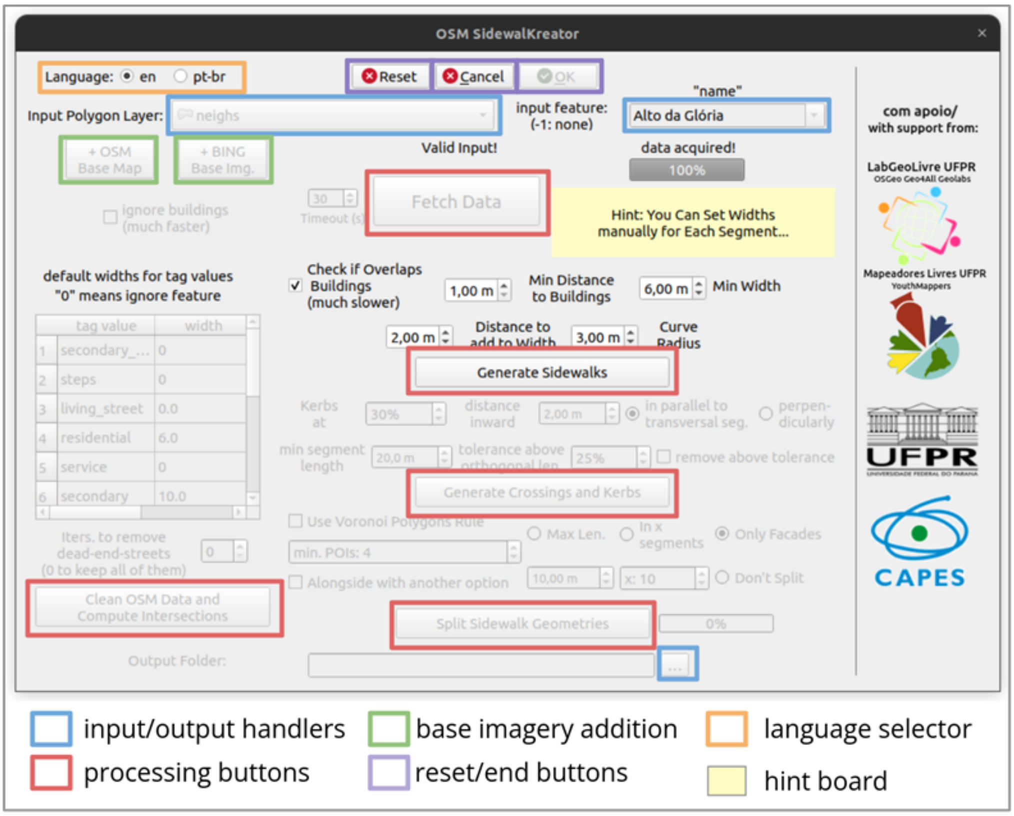

Figure 5. OSM SidewalKreator GUI

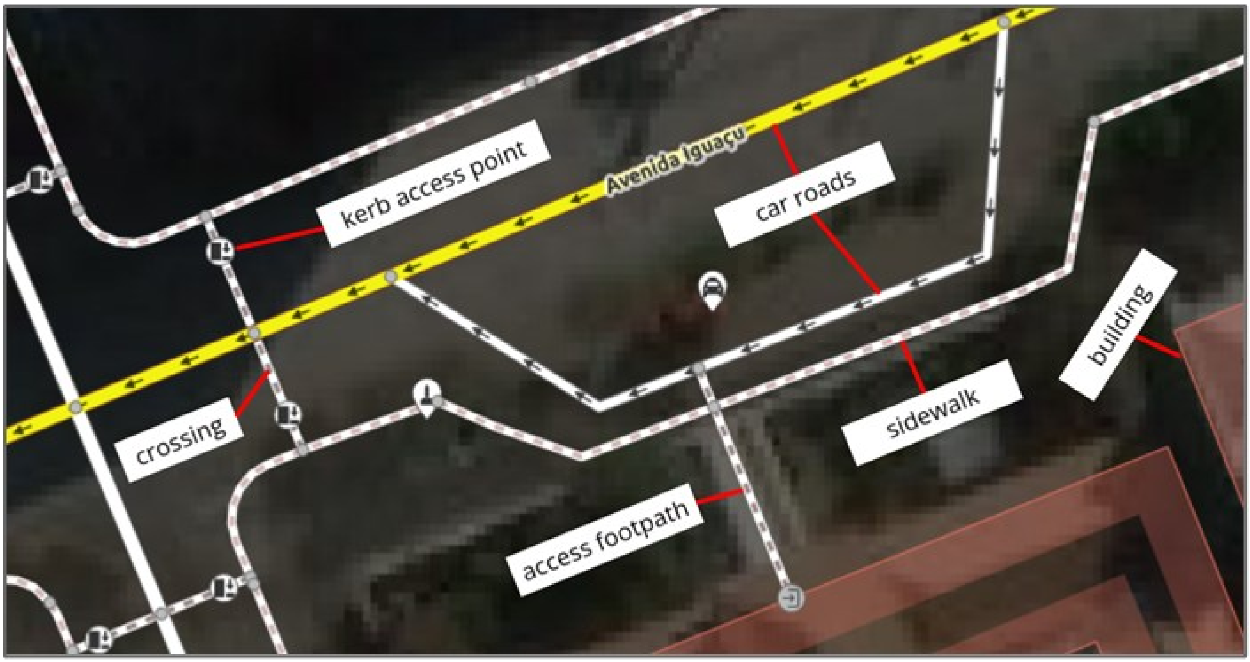

Figure 6. Example of a manually-modified sidewalk geometry

https://doi.org/10.48088/ejg.k.ves.14.4.066.084

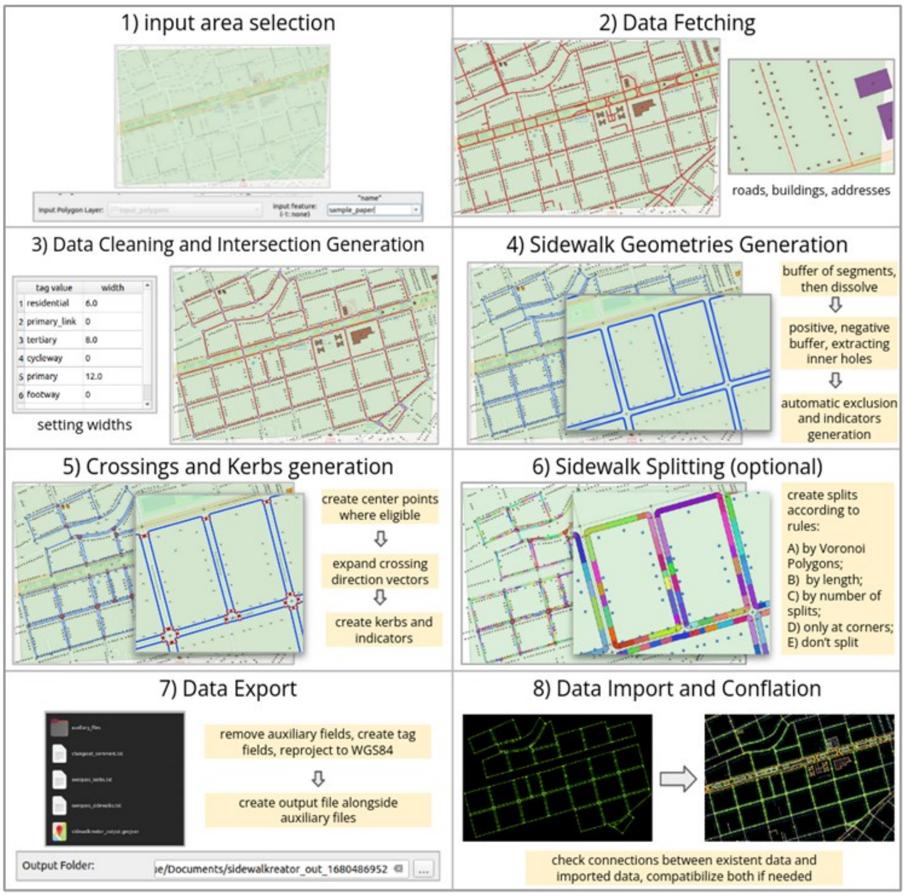

Figure 7. SidewalKreator Procedures Workflow

In steps 4 and 5 of the processing, as previously mentioned there are some user-set parameters that may considerably changes the results. Considering step 4, we generated Figure 8, where we vary discretely two parameters: the curve radius (the user may notice that setting it to zero will create sidewalks with square-corners); and the distance to add to road width (if set to zero the user will be generating the kerb, not the sidewalk).

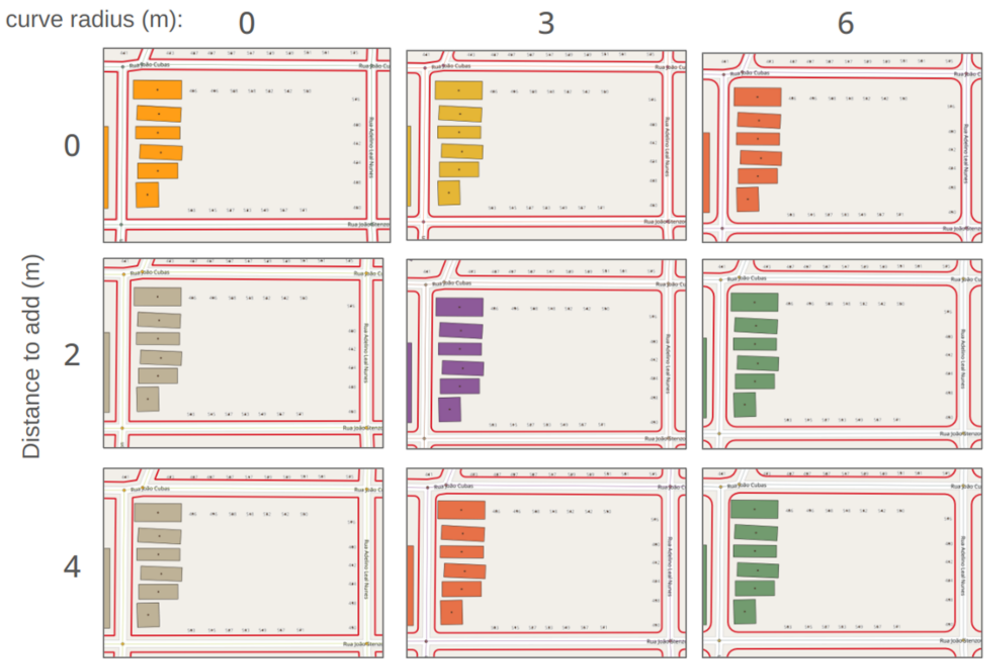

Figure 8. Varying Parameters for Sidewalk Generation

Considering step 5, we generated Figure 9, where we vary discretely two parameters: the distance to add inward for the crossing placement; and the relative (%) placement of kerbs considering half-crossing length. Both got varied in the same way considering the two options for calculating crossing direction: in parallel to nearest adjacent street; and perpendicular to its own street. As on previous figure, the OSM basemap was used to help giving context.

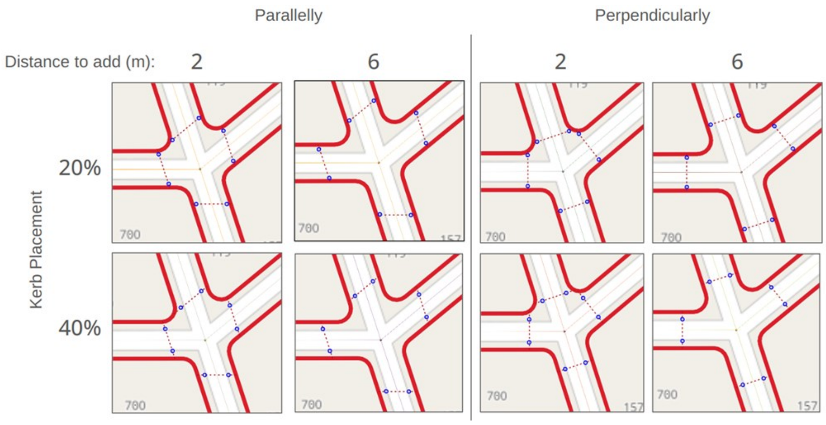

Figure 9. Varying Parameters for Crossings and Kerb Acess Points Generation

In summary, the OSM SidewalKreator plugin aims to be a valuable addition to address the current deficit in OSM sidewalk data. Using an easy-to-use GUI and a comprehensive step-by-step process, the plugin automates the generation of sidewalk geometries and descriptive tags. The plug-in also provides seamless integration with existing OSM data through JOSM, allowing users to merge and update the OSM database with the freshly generated sidewalk information.

5. Evaluation and Discussion

A dedicated GitHub repository has been created to undertake all the analyses in this chapter. It is available on <https://github.com/kauevestena/sidewalk_analysis> in order to enable transparency and reproducibility. Also, there is a configuration file, so it is possible to repeat all those analyses using a different list of neighborhoods.

To evaluate the proposed methodology, a comparative analysis was conducted on the pedestrian network data from two neighborhoods in Curitiba, Água Verde (A) and Jardim das Américas (B). The pedestrian network in Jardim das Américas was manually mapped, while most of the missing pedestrian network features in Água Verde were generated using the SidewalKreator plugin. This comparative approach provides valuable insight into the accuracy and efficiency of the plugin in relation to standard manual mapping techniques. Figure 10 illustrates both neighborhoods with their respective base-map data.

Considering the analysis date, the selected neighborhoods have most of their sidewalks mapped on OSM. They share similar socio-economic characteristics (ABEPRO, 2007) but exhibit geometrical differences in their layouts, as illustrated in Table 1. A focal point in all conducted comparisons is delineating blocks by streets. We refer to these as "proto blocks" because they include simple traffic islands or small squares, and some even lack sidewalks, which may be considered "less significant" within the current context. The proto blocks are derived through the polygonization of line data without applying filters based on attributes such as perimeter or area. Subsequently, this quirk is considered in all forthcoming analyses.

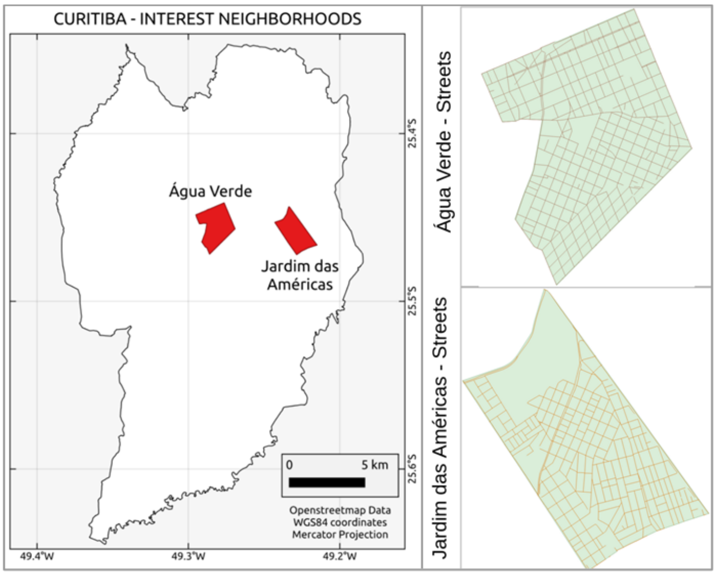

Figure 10. Neighborhoods selected for comparative analysis.

Table 1. Descriptive Attributes of the Compared Neighborhoods

Neighborhood	Água Verde (A)	Jardim das Américas (B)	Difference A-B (%)
OSM Area (m2)	4772117.35	3875653.11	-18.8
Perimeter (m)	9773.68	8847.39	-9.5
Roads Total Length (m)	95073.32	64602.72	-32.0
Protoblocks Count	278	190	-31.7
Protoblocks Mean Area (m2)	17032.19	19857.53	16.6
Protob. Area Std. Dev. (m2)	12773.15	55660.15	335.8
Protoblocks Area Q1 (m2)	11067.96	10658.18	-3.7
Protoblocks Area Median (m2)	16748.62	13718.11	-18.1
Protoblocks Area Q3 (m2)	18974.18	18342.51	-3.3
Protoblocks Min Area (m2)	121461.18	746251.92	514.4
Sidewalks Total Length (m)	137281.62	95711.44	-30.3

It is essential to consider the distinctions between A and B, mainly because B is nearly 20% smaller in area and has 30% fewer sidewalks, roads, and proto blocks drawn. Significant disparities are evident when assessing the variation in the number of proto blocks, mainly due to the presence of the Polytechnic Campus of the Federal University of Paraná in B. Despite spikes in the data, a slight difference is noted in the intermediate quartiles (Q1 and Q3), highlighting the existence of numerous similarly regular proto blocks, as depicted in Figure 10. Nonetheless, substantial differences persist in the average values.

The analyzed pedestrian networks for Água Verde (A) and Jardim das Américas (B) are available for visualization in Figure 11 and Figure 12, respectively. The interested reader will find an interactive counterpart, comprising both A and B, that was made available at https://kauevestena.github.io/sidewalk_analysis/webmap, keep in mind that there are small differences on the rendering side.

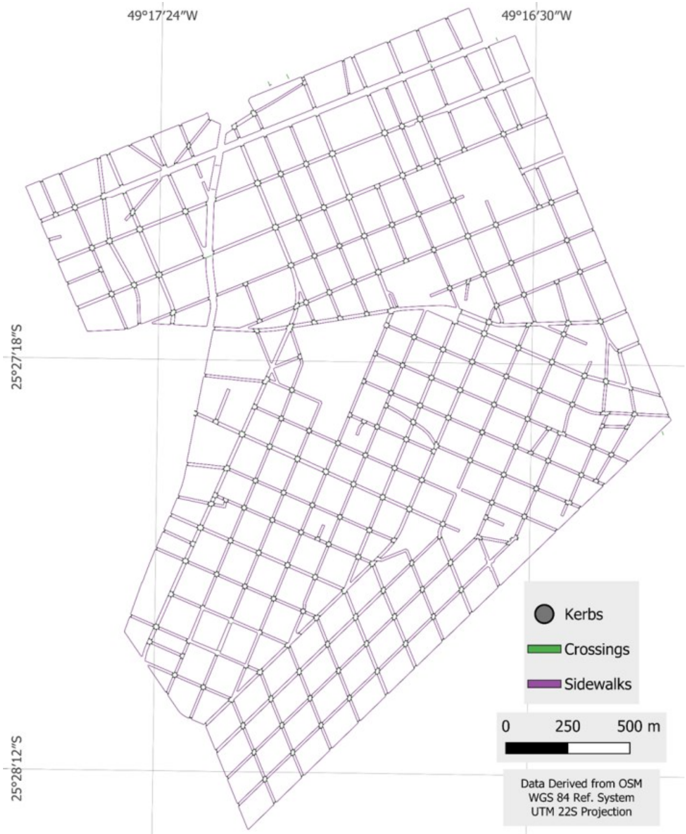

Figure 11. Analysed Pedestrian Network for the Água Verde Neighborhood (A)

https://doi.org/10.48088/ejg.k.ves.14.4.066.084

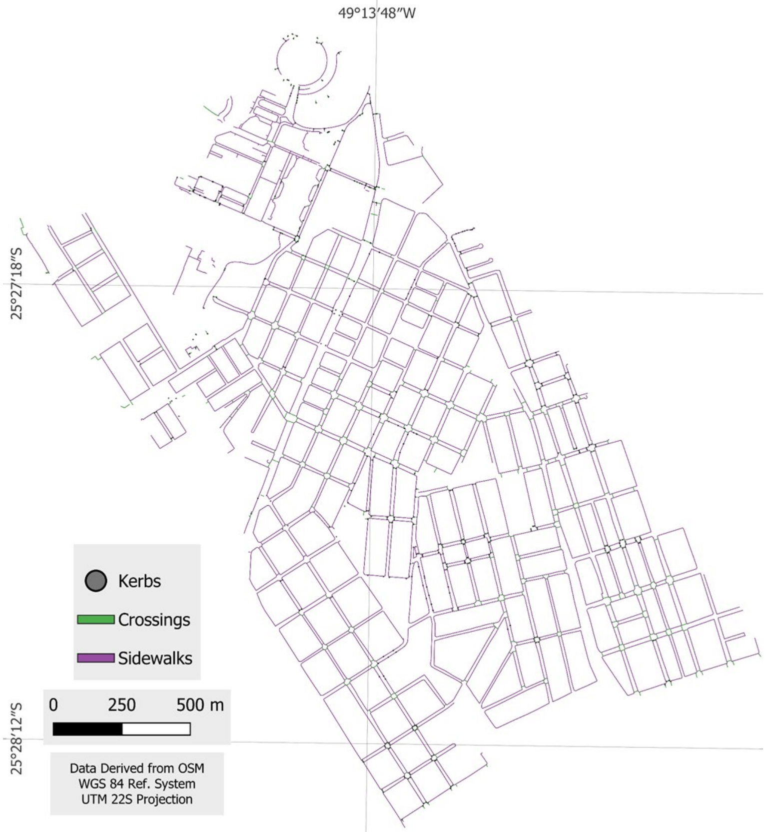

Figure 12. Analysed PedestrianNetwork for the Jardim das Américas Neighborhood (B)

The undertaken analyses are mainly related to data quality concepts. We followed the ISO 19157:2013 (International Organization for Standardization, 2013) standard categories: completeness, thematic accuracy, logical consistency, temporal quality, positional accuracy, and usability. All utilized color palettes are from the ColorBrewer 2.0 tool (Harrower & Brewer, 2003).

The initial analysis focused on the issue of data completeness. In the context of all proto blocks, the simplest, though not universally applicable, condition for assuming completeness of the contained sidewalk features is whether they are 'closed', this was done in a similar way as made by Zhou and Tian (2018), having blocks as the basic unit of analysis for the same goal. This category refers to those features where the start and end points coincide. Table 2 provides a summary of the conditions for each proto-block, including the following conditions: 'Closed Multi', which occurs when the algorithm identifies streets that divide the area under consideration into 'proto blocks', resulting in multiple closed sidewalks within a single 'proto block'; 'Unclosed', which indicates situations where the 'closed' condition cannot be met, typically involving proto blocks with incomplete drawing processes, suggesting potential data issues; and 'Without Sidewalks', which indicates proto blocks with no drawings present, which may include less significant ones, or proto blocks where the drawing process has not yet begun.

Table 2. Classification of proto blocks regarding sidewalk presence

Jardim das Américas		%	Água Verde		%
Closed	143	75.3	Closed	249	89.6
Closed Multi	4	2.1	Closed Multi	1	0.4
Unclosed	12	6.3	Unclosed	2	0.7
Without S.	31	16.3	Without S.	26	9.4

In Figure 13, there is a graphical representation of Table 2. Figure 13 shows that, except for a single case (manually deleted), all 'without sidewalk' instances in A correspond to 'less important' proto blocks. Of the two 'without sidewalk' cases in A, one concerns a traffic island with two intersections erroneously marked as 'sidewalks' (drawn prior to data import), while the other contains an intersection that serves as the only link between two sidewalk points. In B, more than half of the 'without sidewalks' occurrences are regular city blocks where sidewalks exist. There are also 12 instances of 'unclosed sidewalks' with numerous unfinished sidewalk drawings, possibly due to inexperienced mappers.

The next aspect of the plugin's analysis relates to speeding up data generation to improve completeness in more significant scenarios. We conducted the following test: a Python script attempted to replicate the footpath construction process using the same procedure described in section 4, but with the nearest distance to a sidewalk as the buffering distance. We then compared the difference in perimeter between this 'reconstructed sidewalk' and the one available in the OSM database. The central concept is that those in category A with significant perimeter differences are proto blocks that require more extensive preparation for import, often involving manual editing, as described in section 4. This analysis could only be carried out on proto blocks with 'closed' pavements. Figure 14 illustrates the perimeter differences using quartiles.

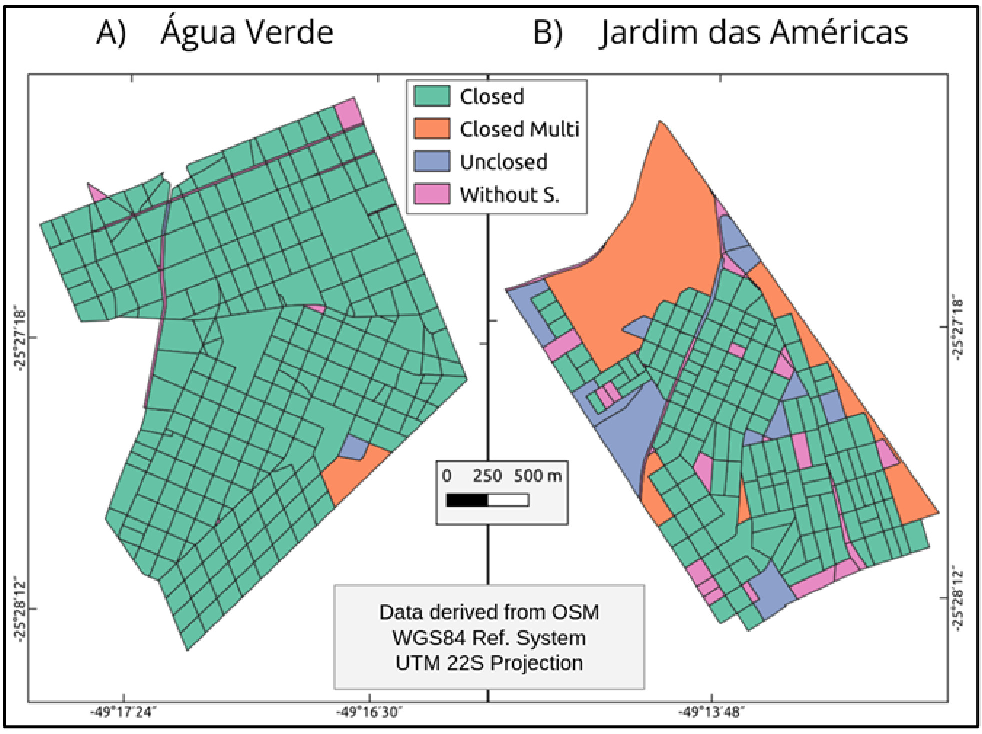

Figure 13. Proto blocks classification regarding sidewalk presence

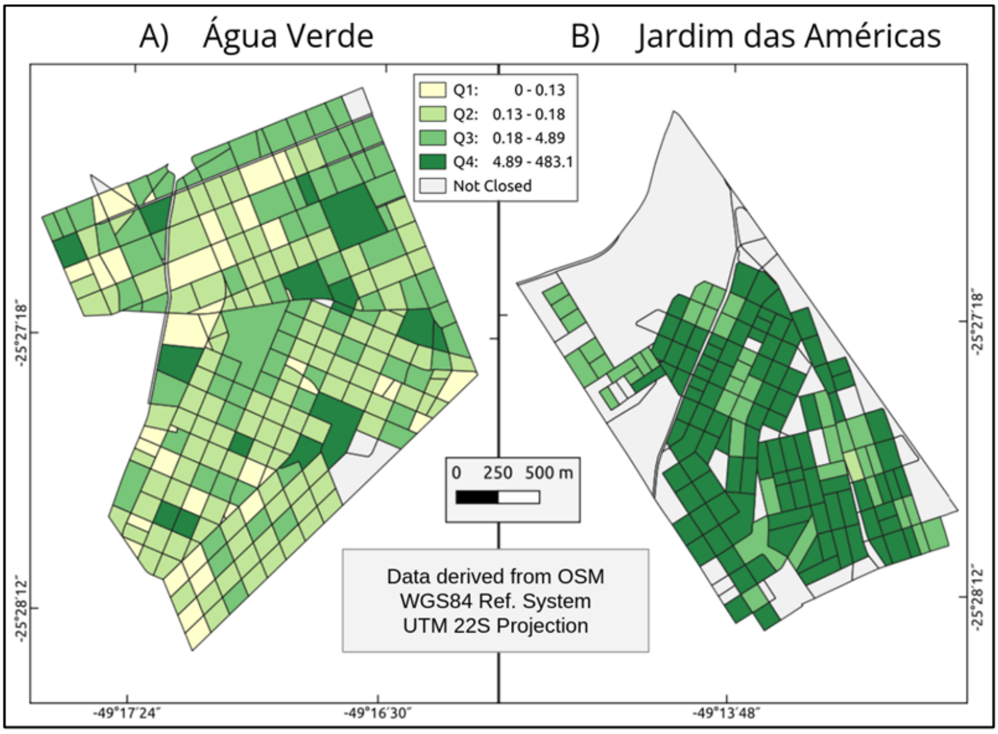

Figure 14. Perimeter differences in quartiles

Readers may observe that there are no samples in Q1 and only a single sample in Q2 for B, with the overwhelming majority falling within Q4. However, it is important to note that the quartiles alone mask the internal variabilities within each neighborhood, as all data was consolidated into a single dataset. To address this, in Figure 15, we created separate box plots for each neighborhood. These box plots are represented with a stacked histogram generated by selecting the data's most significant (in terms of frequency) portion.

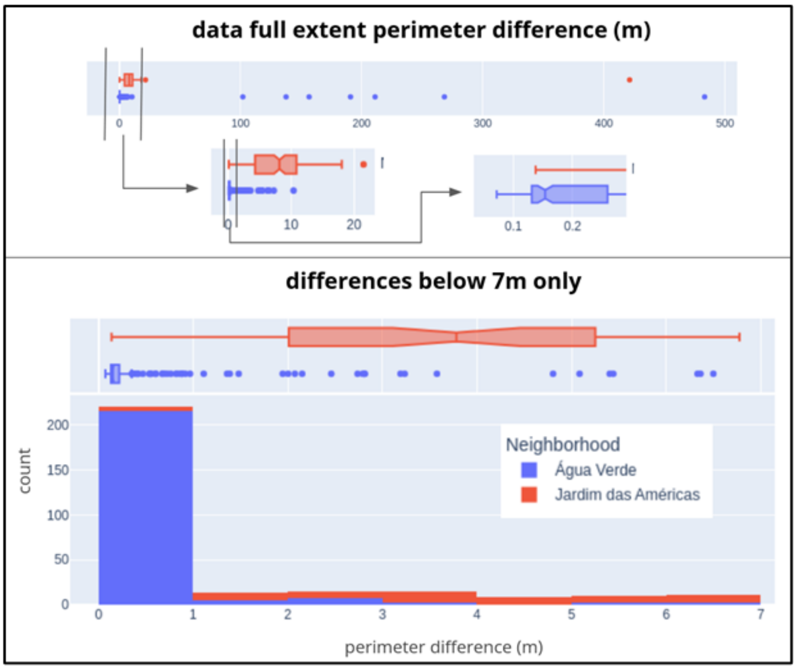

Figure 15. Perimeter differences in quartiles

The boxplots are a valuable tool for highlighting significant variations. In the case of B, the samples have a wide dispersion, indicating a higher degree of irregularity in shape. Approximately 75% of the samples have differences above 4.3 meters, with a median value of 8.18 meters or 3.78 meters for the subset below 7 meters. In A, there are three distinct subgroups: the majority, 78%, have a difference of 1 meter or less (with a Q3 value of 0.25 meters); 19.5% required minor intervention, with differences of up to 10.4 meters; and 2.5% are outliers, with differences of over 100 meters. These outliers correspond to large blocks that either required significant adjustments or more complicated blocks, such as dead-end streets, that the reconstruction algorithm struggled to model because the minimum distance measurement was too small, resulting in non-existent fills, therefore inflated differences.

The next aspect to consider is positional accuracy, which is a challenging aspect of the evaluation. While we did not measure absolute accuracy, we evaluated relative metric accuracy by assuming that smoother paths are generally more accurate (Bodansky & Gribov, 2022). It is worth noting that changes in direction can occur on sidewalk paths, but except for vertices, they tend to exhibit local smoothness, as shown in Figure 16-I. Figure 16-II visually illustrates the smoothness comparison in samples from the tested neighborhoods.

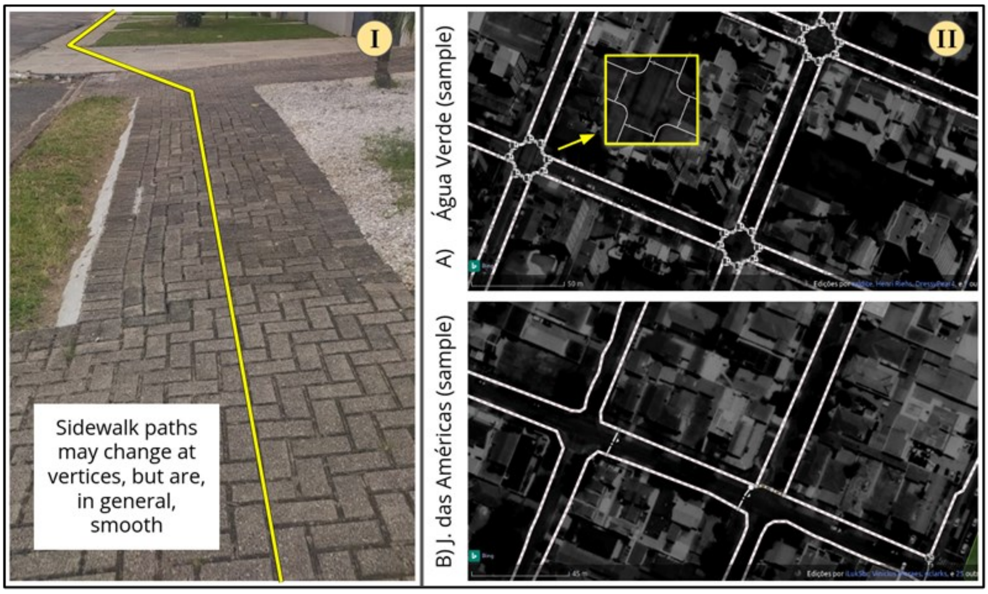

Figure 16. I) example of changing-path sidewalk; II) A x B in smoothness.

In Figure 16-II, the significant difference between A and B is readily apparent through visual inspection. To quantitatively assess this difference in terms of line smoothness, which can be defined in the context of curves by the continuity of their derivatives (Chen, 2010), we calculated the gradient (the quotient of the X and Y coordinate variations) as a discrete approximation of the derivative for each segment. These values have been then averaged using the arithmetic mean to derive a smoothness metric for the entire shape. For this experiment, we selected a sample of 12 simple rectangular blocks where the sidewalk has minimal real-world deviations, thus satisfying the assumptions of smoothness and symmetry. These samples are shown in Figure 17.

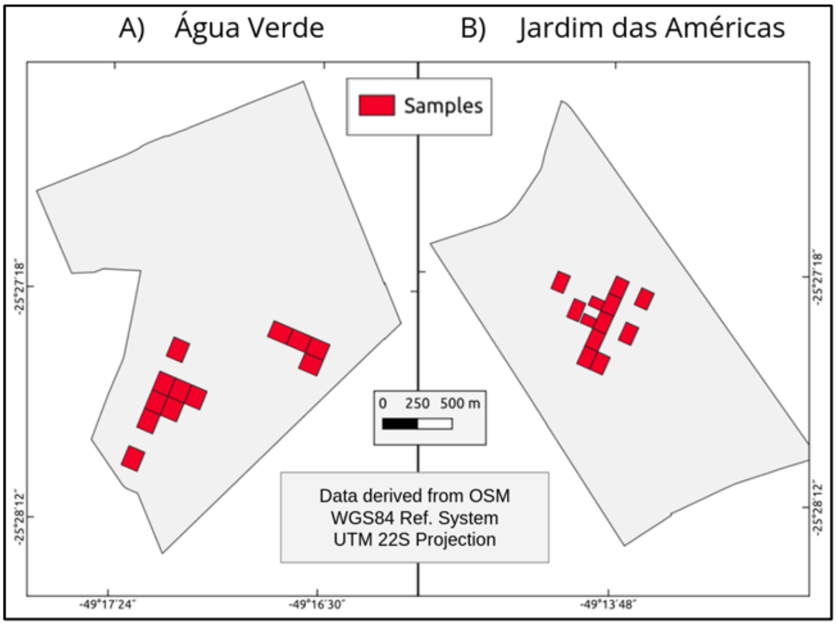

Figure 17. Blocks with sidewalks chosen for shape analysis.

The generated chart comparing the mean of gradients of each of the 12 samples is presented in Figure 18.

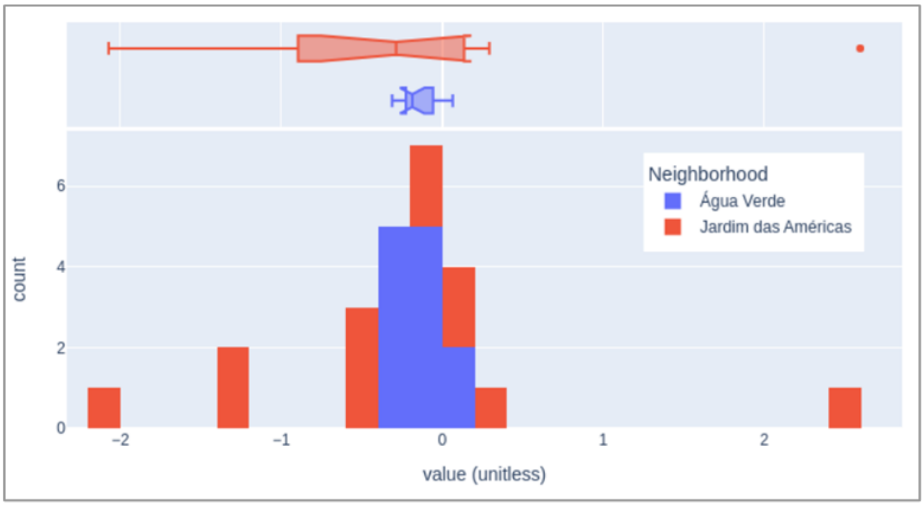

Figure 18. Comparison of Mean Gradients in the samples

In Figure 18, the smaller variability observed in A's samples (ranging from -0.312 to 0.063) indicates the anticipated symmetry present in each algorithmically generated shape. However, within A's set, some asymmetry persists due to differences in rectangular ratios among the samples. In contrast, B's samples exhibit a considerably more comprehensive range (from -2.073 to 2.594), representing a less uniform sample with some individual shapes that are very smooth and symmetrically drawn.

Another metric used to assess relative metric accuracy is the analysis of compactness, which quantifies how much a shape is spread out in a 2D space (Li et al., 2013). Generally, regularly shaped sidewalks tend to be less spread out. For this test, we employed the isoperimetric quotient, also known as "The Polsby-Popper compactness score" (Li et al., 2013), calculated as $4\pi A/P^2$, where A and P represent the area and perimeter of the tested shape, respectively. This score ranges theoretically from 0 (indicating no compactness) to 1, achieved only by a circle – the shape that minimizes the area for a given perimeter and is the most compact. The calculated compactness values for the selected samples are presented in Figure 19-I. In Figure 19-II, in line with the analyses conducted for Figure 14 and Figure 15, we compare the compactness of the sidewalks in OSM DB with their 'reconstructed' counterparts.

https://doi.org/10.48088/ejg.k.ves.14.4.066.084

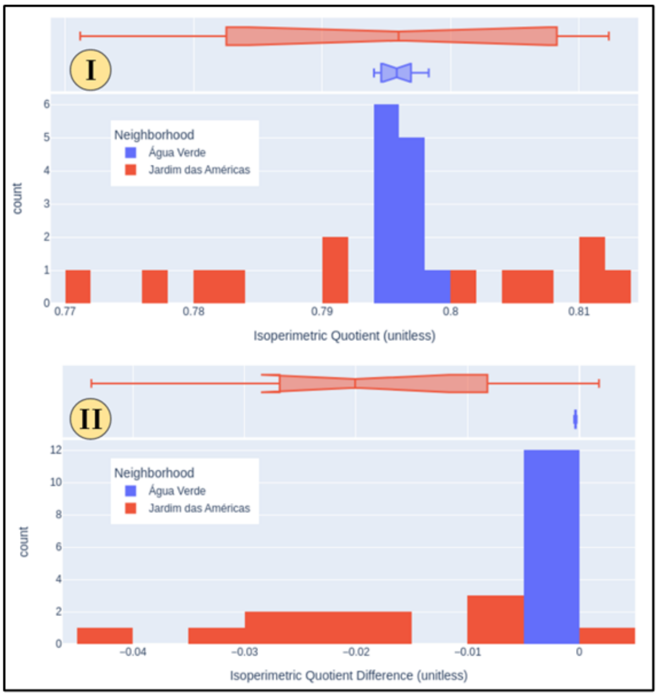

Figure 19. Comparison of Mean Gradients in the samples

Similarly to the phenomena seen in Figure 14 and Figure 15, considering amplitude in Figure 14-I, there is, between A and B sets, a difference of orders of magnitude (42 times bigger), thus showing a much less homogeneous sample for B, which got some samples that are more compact since some were drawn using straight lines, which besides being more compact, do not correspond to the analyzed sample reality, which only contains rectangles with rounded corners. In A, there is a slight asymmetry in the sample, again caused by different axial ratios. Nevertheless, the quotient is closer to 80%, which is not expected to be too close to 100% since the shapes are not circles. Figure 14-II corroborates Figure 14 and Figure 15, with samples in B being way more spread in the 2D space (negative differences) when compared to their automatically drawn counterparts, which are again way less spread. This analysis points out that a human operator has difficulty drawing uniformly regular shapes, a characteristic that is indeed the reality in scenarios like the sampled one.

The subsequent analysis focused on thematic/attribute consistency, as depicted in Figure 13. To delve further into this aspect, we applied a 40% transparency overlay in zoomed portions of Figure 13, resulting in Figure 20. In Figure 20, it becomes apparent that there are paths tagged as 'highway=footway' that lack the 'footway=sidewalk' attribute, which is crucial for completing the logical model and accurately defining the feature as a sidewalk; otherwise, it remains a general footpath, a designation that is not necessarily incorrect but lacks complete precision. This discrepancy is observed in 10 of the 31 proto blocks labeled as 'Without Sidewalks' in Table 2, highlighted in pink as shown in Figure 8.

This inconsistency typically arises from errors made by mappers who may have become distracted during manual drawing, as there is no strict sequence of instructions to follow. Consequently, the incomplete model inadvertently results in a valid OSM feature. An example of a clear, logical error is evident in Figure 16-II, where some crossings incorrectly bear the 'barrier=kerb' tag. This error can be discerned by the rendering style, which differs from that shown in Figure 5.

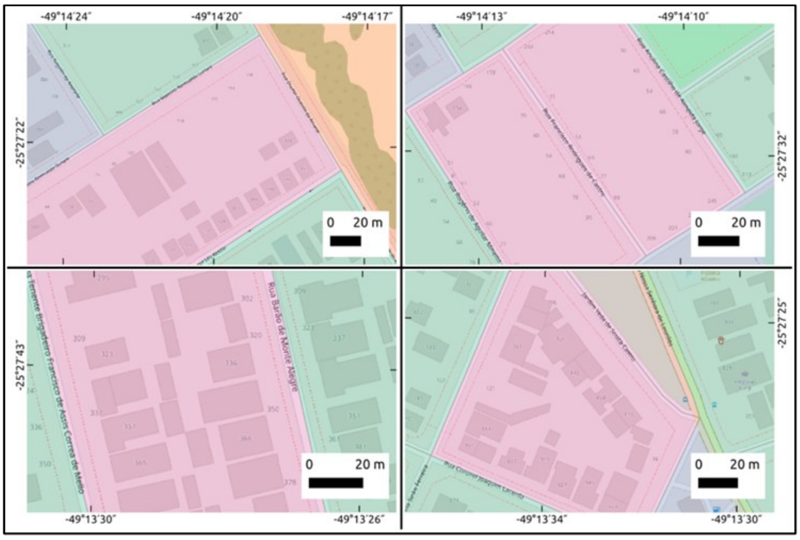

Figure 20. Examples of tag-lacking manually drawn complete sidewalks.

Observing both Figure 20 and Figure 16-II, readers may have noticed the absence of certain crossings, leading us to the next analysis topic. This topic delves into completeness and topological consistency, as crossings play a critical role in ensuring effective connectivity within the network. To explore this phenomenon, we initially examined the total intersecting crossings per 'proto block,' as presented in Figure 21. These crossings are categorized based on what could be interpreted as a 'level of near completeness.' For instance, the expected number of crossings in a simple 4-corner configuration with eight adjoining neighbors is eight.

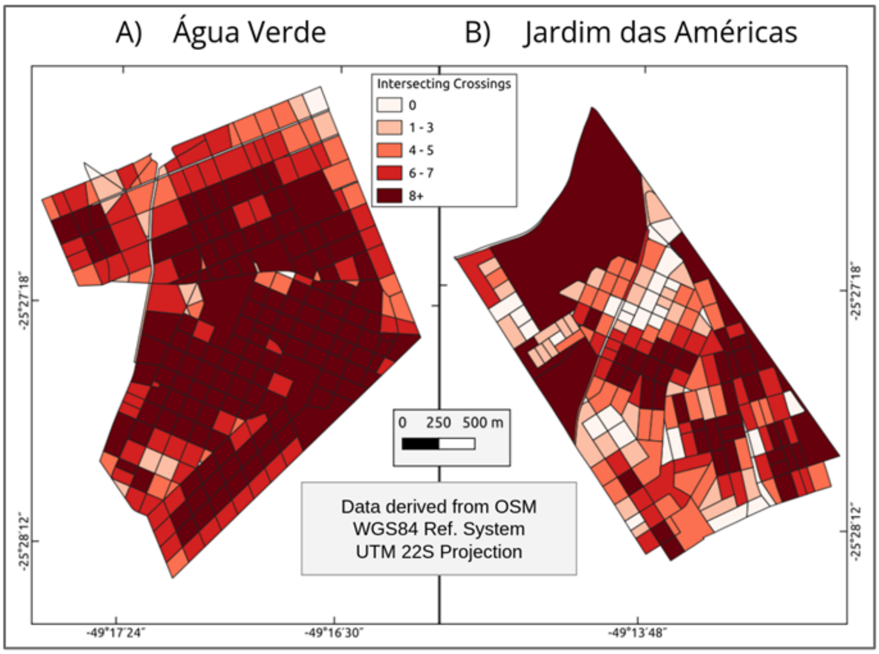

Figure 21. Number of intersecting crossings per proto block

Examining Figure 21, the disparity in completeness in A is unambiguous. In B, as discussed in Figure 13, 25 'important' proto blocks lack a single intersection, rendering their sidewalks, when present, as topological islands. This phenomenon occurs only once in A, while other incomplete cases are more common at the edges. This limitation is due to the current version of the plugin, which restricts the analysis areas to the bounding polygon, excluding proto blocks with sidewalks already drawn. These incompleteness phenomena are particularly visible in two blocks to the southeast of A, and their impact on neighboring areas can be discerned.

Regarding the internal consistency of crossings and kerb access points, we analyzed two characteristic phenomena, presented together in Figure 22: 1) the number of sidewalks each crossing is connected to, ideally 2. A value less than 2 indicates a topological problem, while a value greater than 2 indicates a split at the connection point; 2) the number of kerbs contained within the crossing, which should ideally be two in regular cases. Any value smaller than two indicates a completeness problem.

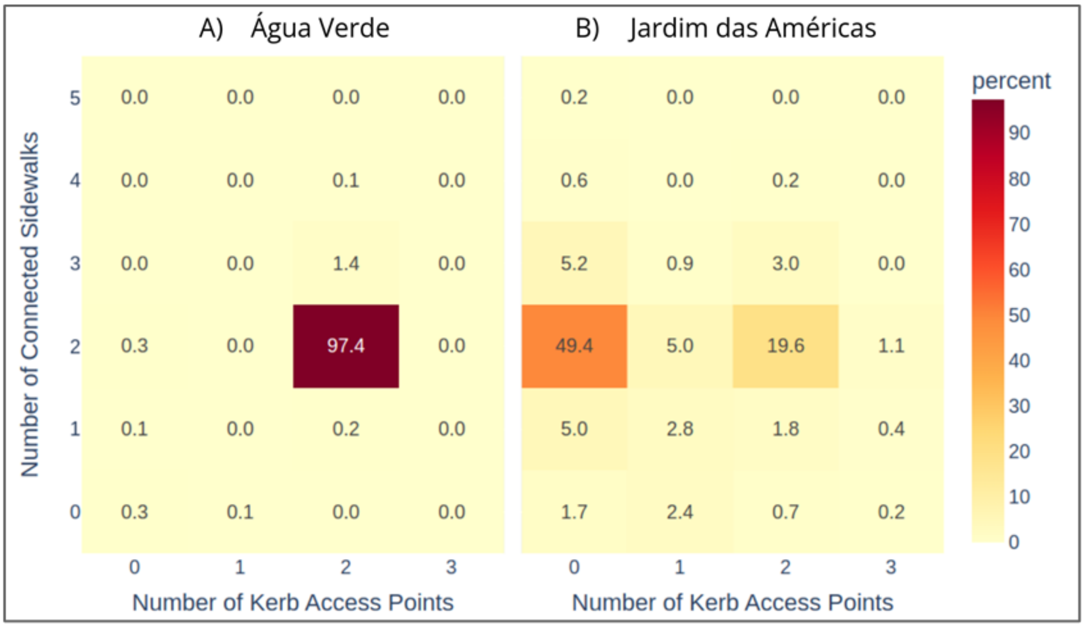

Figure 22. Crossings sidewalk connections and kerb access points quantities

Looking at Figure 22, it is noticeable that the desired combination of standard topology and completeness, characterized by the (x, y) pair (2, 2), is prevalent in A, making up the majority. However, in B this combination represents only one-fifth of the cases, and B also contains 62.1% of the crossings without a single kerb access point, 11.1% with only one kerb access point and even 1.7% with three kerb access points. Regarding sidewalk connections, the B dataset includes 15% with zero or one connection, as opposed to 0.7% in A. It is worth noting that all non-compliant instances in A are residual, mainly due to features created before importing the plugin-generated features, where all crossings are algorithmically constructed to conform to the (2, 2) pair.

Finally, considering data creation distribution in time, we can consider the potential benefits of simultaneously generating all (previously missing) geometries, potentially improving set consistency, because in the neighborhood B manual generation was made by a myriad of users with different levels of dexterity on drawing. Figure 23 illustrates the differences in the distribution of data generation between the two neighborhoods, in A it took a few working hours for an entire neighborhood to be fully available, and in B it took many months for an, in many shown ways, incomplete dataset to be generated.

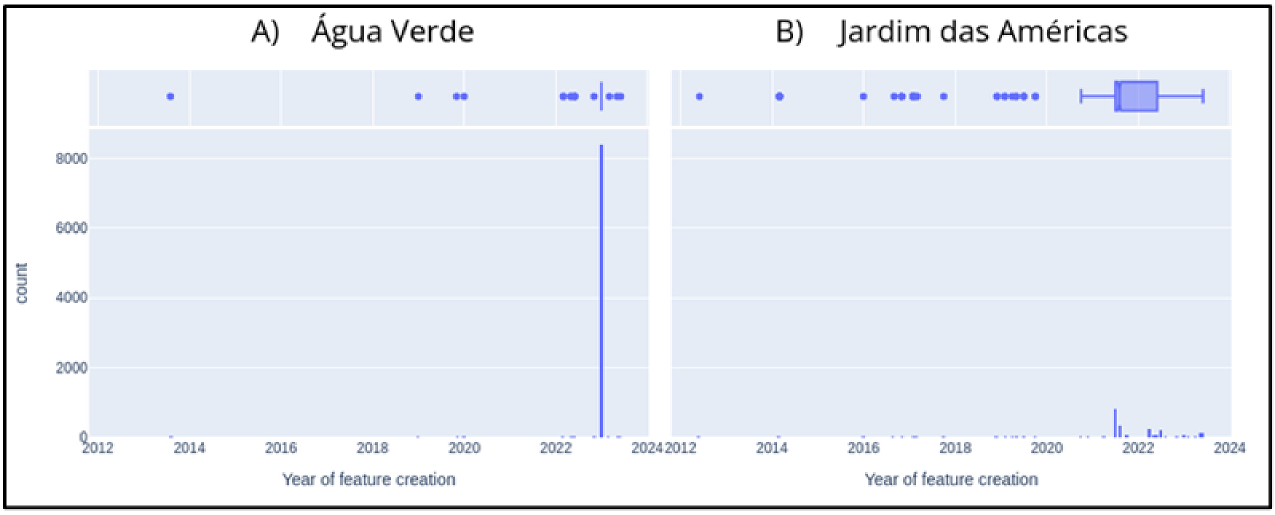

Figure 23. Feature creation timeline

The final analysis pertains to the 'readiness' for future completeness, anticipating the process of populating other feature attributes, such as surface material and smoothness, for example. As discussed in Chapter 4, the optional splitting process generates a larger number of features. This can be particularly advantageous in areas like the city under analysis, where the characteristics of each sidewalk segment along the front of each parcel often undergo frequent changes. In this context, Figure 24 presents the feature density we constructed by generating centroids for all features, considering a blur radius of 150 meters—an empirically derived average block facade size.

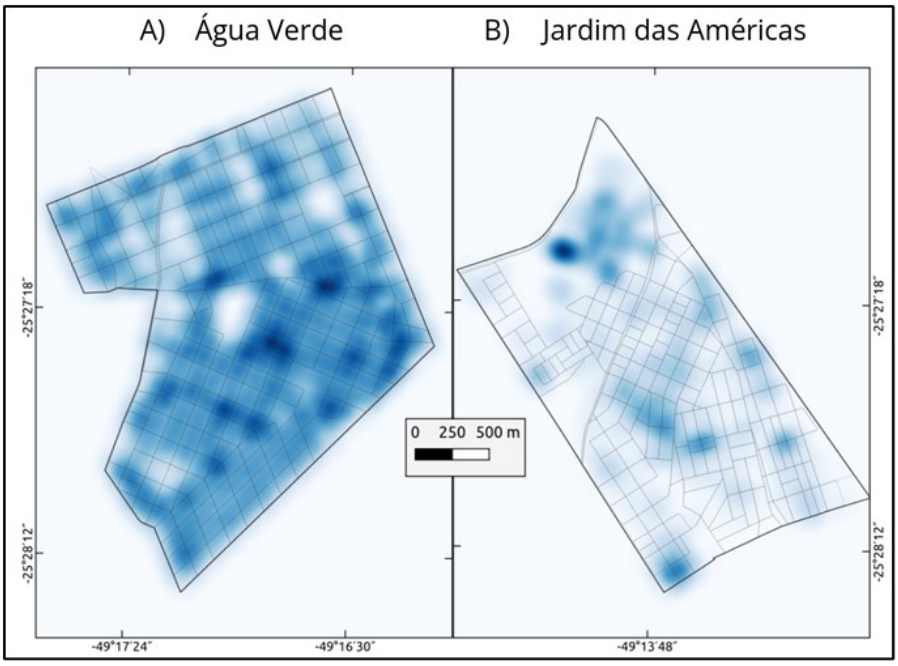

Figure 24. Total feature density per neighborhood

https://doi.org/10.48088/ejg.k.ves.14.4.066.084

The density map once again highlights the disparities in completeness and reveals a significant concentration of non-split sidewalks in B. This concentration may entail a heavier workload in the future, as many sidewalks will likely require splitting to accurately reflect the local non-homogeneous reality.

6. Conclusions

In conclusion, the primary objective of this work is to offer a tool that assists intermediate to advanced GIS and OSM users in expediting the laborious task of manually drawing sidewalks, crossings, and curbs. In this study, we conducted a comprehensive analysis of the OpenSidewalkKreator plugin, aiming to assess its effectiveness in enhancing data completeness, accuracy, and topological consistency within two distinct urban neighborhoods—Area A, where the plugin was employed, and Area B, where sidewalks were manually added. Our findings reveal several vital insights that underscore the plugin's potential impact on urban data quality and its comparison to manual data entry. As long as there is no automated way of getting essential parameters and due to the demand for manual steps in parts of the conflation, the SidewalKreator plugin can't be deemed as a fully automated one-click solution, rather it emphasizes the importance of active user involvement in reviewing results step-by-step, harnessing the capabilities of both QGIS and JOSM toolsets in order to achieve optimal results. After importing the data, users are encouraged to undertake a Tasking Manager validation project to ensure accuracy.

The main limitations of the method include its inability to handle multi-level roads, such as overpasses, the frequent lack of accurate information on OSM street widths, and the uncertainty of the very presence of a sidewalk. In addition, the lack of integration with satellite and/or terrestrial imagery could be addressed in future iterations of the plugin, to cover up for example the detection of the presence of crossing restrictions or sidewalk inexistence. There are also plans to extend the functionality of the plugin to make it available as an easily integrated QGIS processing provider algorithm, with additional algorithms such as drawing missing intersections and access points. Despite its current limitations, the SidewalKreator plugin has the potential to be a valuable tool for improving the representation of sidewalks in OSM.

While advocating for using separate geometries as the optimal way to represent sidewalk networks, this work also recognizes the value of the existing tag schema in specific situations, constantly reminding the user that decisions from local communities prevail regarding choosing concurrent schemas.

The undertaken analyses pointed out many fragilities of the mapping carried out by human operators, some caused by the difficulties in drawing regular shapes and keeping the needed homogeneity to correct model the reality and comply with the entire data model, which needs thematic and topological consistency.

By combining increased awareness within mapping communities, detailed guidelines, and automation tools like the OSM SidewalKreator plugin, we can work towards a more pedestrian-centered approach to mapping. This will ultimately support applications such as optimized pedestrian routing, including tailored profiles for individuals with disabilities. In the long term, the outcomes of this project aim to establish a robust foundation for enhancing accessibility and mobility for cities worldwide.

Supplementary Material: All the supporting materials are available on two GitHub repositories: https://github.com/kauevestena/osm_sidewalkreator corresponding to the plugin source code and direct additional information; and https://github.com/kauevestena/sidewalk_analysis containing all the scripts for the analyses carried out for this publication.

Funding: This study was financed in part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance Code 001.

Acknowledgments: For the *Mapeadores* Livres group, part of the YouthMappers network and all OpenStreetMap community, from where the manual mappers came, thus giving the manual data for our comparisons.

Conflicts of Interest: The authors declare no conflict of interest.

References

ABEPRO. (2007). Indicadores de qualidade do ambiente para Curitiba: Uma proposta de modelo. Retrieved from https://abepro.org.br/biblioteca/enegep2007_tr680486_0573.pdf

Bartzokas-Tsiompras, A., & Photis, Y. (2020). Does neighborhood walkability affect ethnic diversity in Berlin? Insights from a spatial modelling approach. *European Journal of Geography*, 11(1), 163–187. https://doi.org/10.48088/ejg.a.bar.11.1.163.18

Bartzokas-Tsiompras, A., Photis, Y. N., Tsagkis, P., & Panagiotopoulos, G. (2021). Microscale walkability indicators for fifty-nine European central urban areas: An open-access tabular dataset and a geospatial web-based platform. *Data in Brief*, 36, 107048. https://doi.org/10.1016/j.dib.2021.107048

Bartzokas-Tsiompras, A., Bakogiannis, E., & Nikitas, A. (2023). Global microscale walkability ratings and rankings: A novel composite indicator for 59 European city centres. *Journal of Transport Geography*, 111, 103645. https://doi.org/10.1016/j.jtrangeo.2023.103645

Biagi, L., Brovelli, M. A., & Stucchi, L. (2020). Mapping The Accessibility In Openstreetmap: A Comparison Of Different Techniques. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B4-2020*, 229–236. https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-229-2020

Bodansky, E., Gribov, A., & Pilouk, M. (2002). Smoothing and Compression of Lines Obtained by Raster-to-Vector Conversion. *Graphics Recognition Algorithms and Applications*, 256–265. https://doi.org/10.1007/3-540-45868-9_22

Bustos, C., Rhoads, D., Solé-Ribalta, A., Masip, D., Arenas, A., Lapedriza, A., & Borge-Holthoefer, J. (2021). Explainable, automated urban interventions to improve pedestrian and vehicle safety. *Transportation Research Part C: Emerging Technologies*, 125, 103018. https://doi.org/10.1016/j.trc.2021.103018

Camboim, S., Bravo, J., & Sluter, C. (2015). An investigation into the completeness of, and the updates to, OpenStreetMap data in a heterogeneous area in Brazil. *ISPRS International Journal of Geo-Information*, 4(3), 1366–1388. https://doi.org/10.3390/ijgi4031366

Chen, L. (2010). Smoothness and smooth extensions (I): Generalization of MWK functions and gradually varied functions. arXiv preprint https://doi.org/10.48550/arXiv.1005.3727

Ertz, O., Fischer, A., Ghorbel, H., Hüsser, O., Sandoz, R., & Scius-Bertrand, A. (2021). Citizen Participation & Digital Tools To Improve Pedestrian Mobility In Cities. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-4/W1, 29–34. https://doi.org/10.5194/isprs-archives-XLVI-4-W1-2021-29-2021

Ferster, C., Fischer, J., Manaugh, K., Nelson, T., & Winters, M. (2019). Using OpenStreetMap to inventory bicycle infrastructure: A comparison with open data from cities. *International Journal of Sustainable Transportation*, 14(1), 64–73. https://doi.org/10.1080/15568318.2018.1519746

Fidalgo, C. D., Machado Santos, I., de A. Nogueira, C., Clara Souza Portugal, M., & Maria T. Martins, L. (2021). Urban sidewalks, dysfunction and chaos on the projected floor. The search for accessible pavements and sustainable mobility. In Proceedings of the 7th International Congress on Scientific Knowledge. https://doi.org/10.25242/8876113220212366

Froehlich, J. E., Saugstad, M., Saha, M., & Johnson, M. (2022). Towards Mapping and Assessing Sidewalk Accessibility Across Sociocultural and Geographic Contexts. https://doi.org/10.48550/arXiv.2207.13626

Gjeruldsen, E. (2020). Detecting sidewalks in OpenStreetMap using machine learning LATEX Version. Master's thesis, Østfold University. https://hdl.handle.net/11250/2676606

Haklay, M. (2010). How good is OpenStreetMap information? A comparative study of OpenStreetMap and Ordnance Survey datasets for London and the rest of England. *Environment and Planning B: Planning and Design*, 37, 682-703. https://doi.org/10.1068/b35097

Harrower, M., & Brewer, C. A. (2003). ColorBrewer. org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1), 27-37. https://doi.org/10.1179/000870403235002042

Hochmair, H. H., Zielstra, D., & Neis, P. (2014). Assessing the Completeness of Bicycle Trail and Lane Features in OpenStreetMap for the United States. *Transactions in GIS*, 19(1), 63–81. https://doi.org/10.1111/tgis.12081

Hosseini, M., Saugstad, M., Miranda, F., Sevtsuk, A., Silva, C. T., & Froehlich, J. E. (2022). Towards Global-Scale Crowd+AI Techniques to Map and Assess Sidewalks for People with Disabilities. https://doi.org/10.48550/arXiv.2206.13677

International Organization for Standardization. (2013). Geographic information — Data quality (ISO Standard No. 19157:2013). https://www.iso.org/standard/32575.html

Jiang, Y., Lobo, M. J., Christophe, S., & Jouffrais, C. (2021). Mapping road crossings for visually impaired people. In 30th International Cartographic Conference (ICC 2021), Florence, Italy, 14-18 December 2021 (pp. 1–2). https://doi.org/10.5194/ica-abs-3-127-2021

Kayama, R., & Yairi, I. (2013). Detection of sidewalk border using camera on buggy for people with visual impairments. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 17(1), 114-119. https://www.semanticscholar.org/paper/Detection-of-sidewalk-border-using-camera-on-buggy-Kayama-Yairi/416337071e9babd9f9193e17c42f607add54e4af

Lei, T. L. (2020). Geospatial data conflation: a formal approach based on optimization and relational databases. *International Journal of Geographical Information Science*, 34(11), 2296–2334. https://doi.org/10.1080/13658816.2020.1778001

Leśniewski, M., Świąder, M., Kaczmarek, I., Castro, D., Kamińska, J. A., Pilawka, T., & Kazak, J. K. (2021). Towards Transit-Oriented Development for Sustainable Urban Mobility: Insights from a Central European City. *Geomatics and Environmental Engineering*, 15, 39-53. https://doi.org/10.7494/geom.2021.15.3.39

Li, W., Goodchild, M. F., & Church, R. (2013). An efficient measure of compactness for two-dimensional shapes and its application in regionalization problems. *International Journal of Geographical Information Science*, 27(6), 1227-1250. https://doi.org/10.1080/13658816.2012.752093

Map Features. (2023). In OpenStreetMap Wiki. Retrieved March 8, 2023, from https://wiki.openstreetmap.org/wiki/Map_features

Mobasheri, A., Bakillah, M., Rousell, A., Hahmann, S., & Zipf, A. (2015). On the completeness of sidewalk information in OpenStreetMap, a case study of Germany. In Proceedings of the 18th AGILE International Conference on Geographic Information Science, Lisbon, Portugal, 9-12 June 2015. https://api.semanticscholar.org/CorpusID:4483034

Mobasheri, A., Huang, H., Degrossi, L., & Zipf, A. (2018). Enrichment of OpenStreetMap data completeness with sidewalk geometries using data mining techniques. *Sensors*, 18(2), 509. https://doi.org/10.3390/s18020509

Mobasheri, A., Sun, Y., Loos, L., & Ali, A. (2017). Are crowdsourced datasets suitable for specialized routing services? Case study of OpenStreetMap for routing of people with limited mobility. *Sustainability*, 9(6), 997. https://doi.org/10.3390/su9060997

Mooney, P., & Minghini, M. (2017). A Review of OpenStreetMap Data. In Mapping and the Citizen Sensor (pp. 37–59). Ubiquity Press. https://doi.org/10.5334/bbf.c

Neis, P., & Zielstra, D. (2014). Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future Internet*, 6(1), 76-106. https://doi.org/10.3390/fi6010076

Neis, P., & Zipf, A. (2012). Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap. *ISPRS International Journal of Geo-Information*, 1(2), 146–165. https://doi.org/10.3390/ijgi1020146

New York City Department of Transportation. (2023). About DOT. Retrieved from https://www.nyc.gov/html/dot/html/about/about.shtml

Omar, K. S., Moreira, G., Hodczak, D., Hosseini, M., & Miranda, F. (2022). Crowdsourcing and Sidewalk Data: A Preliminary Study on the Trustworthiness of OpenStreetMap Data in the US. https://doi.org/10.48550/arXiv.2210.02350

OSM contributors. (2023). Taginfo. https://taginfo.openstreetmap.org

OSM Wiki. (2011). Proposed features/Sidewalk as separate way. Retrieved from https://wiki.openstreetmap.org/w/index.php?oldid=2312766

OSM Wiki. (2016). Proposed features/sidewalk schema. Retrieved April from https://wiki.openstreetmap.org/wiki/Proposed_features/sidewalk_schema

OSM Germany. (2023). JOSM: an advanced and extensible OSM editor. https://josm.openstreetmap.de/wiki/WikiStart

Pereira, K. C., Neckel, A., Kujawa, H. A., Kunz, M., Cardoso, G., Ribeiro, L. A., de Almeida Silva, C. C. O., Gnoatto, P., Frighetto, J., & Visentin, T. G. (2017). Pedestrian Mobility on Public Sidewalks Evaluated by the IAAPE Method. *Journal of Civil Engineering and Architecture*, 11(8). https://doi.org/10.17265/1934-7359/2017.08.007

Bustos, C., Rhoads, D., Solé-Ribalta, A., Masip, D., Arenas, A., Lapedriza, A., & Borge-Holthoefer, J. (2021). Explainable, automated urban interventions to improve pedestrian and vehicle safety. *Transportation Research Part C: Emerging Technologies*, 125, 103018. https://doi.org/10.1016/j.trc.2021.103018

Rodríguez, D. A., Merlin, L., Prato, C. G., Conway, T. L., Cohen, D., Elder, J. P., Evenson, K. R., McKenzie, T. L., Pickrel, J. L., & Veblen-Mortenson, S. (2015). Influence of the built environment on pedestrian route choices of adolescent girls. *Environment and behavior*, 47(4), 359–394. https://doi.org/10.1177/0013916513520004

Saha, M., Saugstad, M., Maddali, H. T., Zeng, A., Holland, R., Bower, S., Dash, A., Chen, S., Li, A., Hara, K., & Froehlich, J. (2019). Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3290605.3300292
Sehra, S., Singh, J., & Singh Rai, H. (2013). Assessment of OpenStreetMap Data - A Review. *International Journal of Computer Applications*, 76(16), 17–20. https://doi.org/10.5120/13331-0888
Sidewalk city: remapping public space in Ho Chi Minh City. (2015). *Choice Reviews Online*, 53(02), 53-0613-53–0613. https://doi.org/10.5860/choice.192737
Taskar Center for Accessible Technology & Humanitarian Openstreetmap Team. (2021). OpenSidewalks Mapping statistics Vila Mariana, Sao Paulo, Brazil - Crossings. Retrieved from https://tasks.opensidewalks.com/projects/51/stats
Taskar Center for Accessible Technology & Humanitarian OpenStreetMap Team. (2021). OpenSidewalks Tasking Manager. Retrieved from https://tasks.opensidewalks.com
Taskar Center for Accessible Technology, University of Washington. (2016). OpenSidewalks. Retrieved from https://www.opensidewalks.com
UN. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf
US Department of Transportation, Federal Highway Administration - FHWA Safety Program. (2013). Safety benefits of walkways, sidewalks, and paved shoulders. https://safety.fhwa.dot.gov/ped_bike/tools_solve/walkways_brochure/walkways_brochure.pdf
Vestena, K. M. (2021). OSM SidewalKreator. https://github.com/kauevestena/osm_sidewalkreator
Zhou, Q., & Tian, Y. (2018). The use of geometric indicators to estimate the quantitative completeness of street blocks in OpenStreetMap. *Transactions in GIS*, *22*(6), 1550–1572. https://doi.org/10.1111/tgis.12486

European Journal of Geography