Project Name:

OpenSidewalks Data Schema, and validation tools

### **Version**: 0.1

**Last Modified**: May 25th, 2020

**Client**: Taskar Center for Accessible Technology

**Point of Contact**: Dr. Anat Caspi

-

# 

[**Data Schema Requirements**](#_aanz0wpd7ogf) **2**

[**Validation Requirements**](#_rmaiyaml7tis) **2**

[**Clean up Reqs**](#_c4qzza17u3s3) **3**

[**Testing DAta**](#_q1ki9n3ub3uv) **3**

[**Authentication**](#_67pc5nohckp2) **3**

[**Identification Requirements**](#_9maetmpludcb) **3**

# 

# Data Schema Requirements

1. Read, and understand the current data schema from <https://github.com/OpenSidewalks/OpenSidewalks-Schema>
2. Work with the client to enhance the data schema
3. Publish the updated data schema on github pages

# Validation Requirements

1. Given an input, validate:
   1. Required tags from the data schema exist for each sidewalk segment
2. Generate:
   1. validated.geojson: output of all the segments that adhere to the data schema
   2. split.geojson: file showing the information of the sidewalk segments we would like to split for better routing
   3. unrecognized.geojson: file with sidewalks with no recognized tags from the data schema
   4. private.geojson: file with sidewalks that have required data schema tag, but also have additional information
   5. valid\_routes.geojson: file with the validated segments, and the full connected graph
   6. private\_valid\_routes.geojson: file with validated, and private segments with full connectivity
3. Input could be
   1. a bounding box from OSM
   2. A presumably adherent file

# Clean up Reqs

1. Join the nodes that are “Really close”. Definition of “really close” needs to be determined
2. Split the segments as needed for routing
3. Add a missing identifier
4. Additionally, an API should be provided to take two geojson files, and combine them into one unified geojson file

# Testing Data

1. MS Redmond data - even though it is not exactly an open set
2. Austin data
3. San Jose
4. NYC - Sidewalk labs data

# Authentication

1. Each client needs to obtain an API Key that is needed to make an API call

# Identification Requirements

For each validated sidewalk segment:

1. Add an identifier, tying it with the identifier from shared street identifier
   1. Exact ID specification needs to be worked out. This ID specification should also be included in data schema documentation

Tasks:

1. Add a reference ID to the sidewalks. Sidewalks = (set of nodes, and edges)
2. Anat - push the SSID

Notes:

* Take projection, may have to split the sidewalk segments into multiple depending on the attributes

Usecases:

* OSM -> OSW data mapping
* Point out the conflicts from OSW mapped to the given reference geojson file.
  + Parameterizable “closeness”. 1m as the main set
  + Highlight the differences
* Filter out some of the tags, or filter in only some tags (priority: ?? )

# Elements Validation Requirements

### Node

* Standalone nodes or disconnected nodes:
  + These can only be points like trees, benches etc,
  + Tags are a must for such cases
* Unless otherwise specified, all nodes must be part of ways
  + Tags are optional in this case
* All “kerbs” are nodes
* If two nodes are <1cm apart, consider them the same nodes (from accessmap code base)

### Ways

* at least one primary highway tag.
* Calculate length of the ways (provide an API for this)

### Questions:

* How are the nodes in a way sorted? Some seem to go from S->N, some W->E, and some from E->W. Is it arbitrary?
  + Right now it does, but we could impose a W->E, and S->N convention
* How are IDs generated in OSM? When someone gives us geojson file, is it possible that the node IDs conflict with already existing IDs in OSM?
* Are we interested in pedestrian streets, and biking path,? Yes
* Why is length part of the “ways”? It can be calculated on the fly, but is kept as part of the standard. This may not be necessary to be edited by the “inputter” but can be made part of the intake.
* What is the “foot” keyword?
  + Who has the right of way
* Brunnel?
* “Incline” ? min, max,?
  + Should incline be part of the data schema? Or is it part of AccessMap data?
* Is the OSW sidewalk data sidewalk centerline? Or a polygon structure?
* Use this as example: <https://www.openstreetmap.org/way/477389505>
* What are sidewalk-bulbs?
* Suggestions on potential connectivity:
  + One edge link with the existing nodes. For example, sidewalk corner to the ramp, as a crosswalk/sidewalk\_link
* Rollability:
  + Is there a rollable path?
  + Marked path?
  + Shortest distance?
* If there is a crossing, it needs to be connected to a sidewalk\_link
* If there is a crossing, and a nearby kerb\_ramp, it can be connected.

|  |
| --- |
| Pick the picture that best matches the markings of this crosswalk.  Screen%20Shot%202017-10-19%20at%2011.59.46%20AM.pngScreen%20Shot%202017-10-19%20at%2012.03.53%20PM.png  Screen%20Shot%202017-10-19%20at%2012.08.05%20PM.png  Screen%20Shot%202017-10-19%20at%2012.08.27%20PM.png  download.png |

## Use Cases

### Open Columbus

These folks collected

https://github.com/openmobilityfoundation/mobility-data-specification