OpenStreetMap: Quality, Usage, Verification

Drafted by Trillium Solutions for Northwest Oregon Transit's website trip planner enhancement

December2020

nwCONNECTOR
PUBLIC TRANSIT ACROSS NW OREGON

Trillium
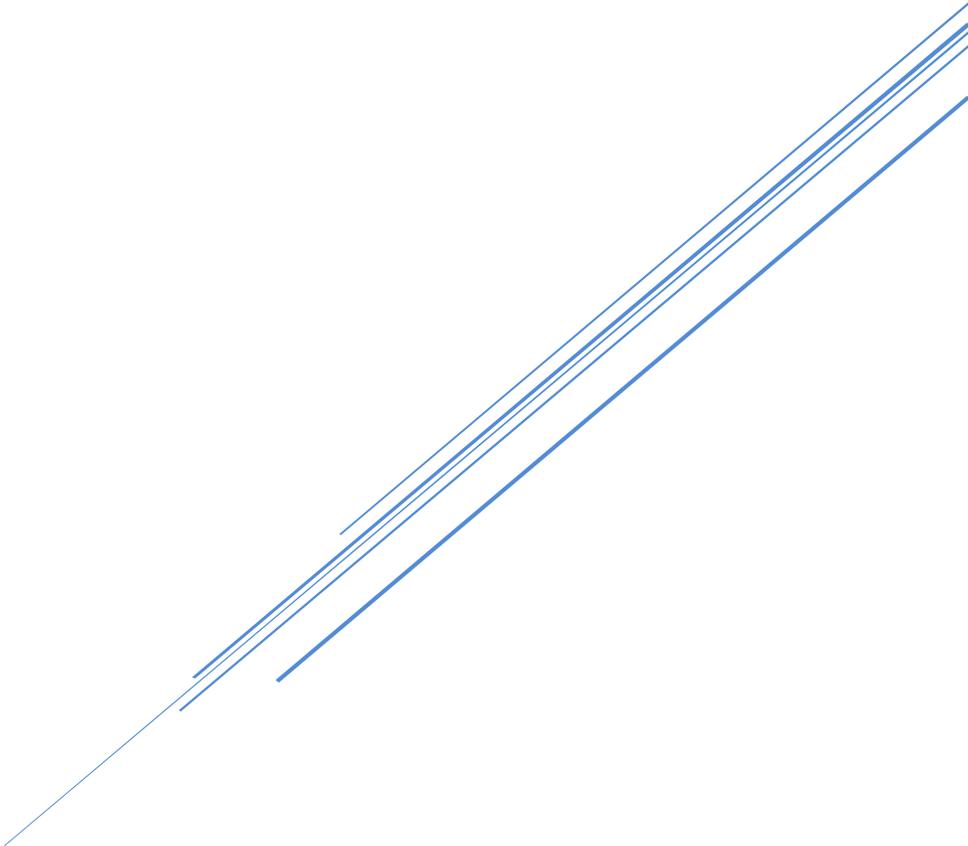
Table of Contents

What is OpenStreetMap?

OpenStreetMap (OSM) is a free, editable map of the world that is continually being built and edited by volunteers. All data in OpenStreetMap have either been donated or created by people like you and me. Based on a "wiki" model (think Wikipedia), anyone can contribute, and edits are logged for accountability. OpenStreetMap is licensed under the Open Data Commons Open Database License (ODbL), which means that anyone is free to use it as long as OpenStreetMap is credited and derivative works are released under the same license.

What is OpenTripPlanner?

OpenTripPlanner is an open source trip planner that can generate itineraries combining multiple modes including transit, pedestrian, bicycle, car, and more recently, demand response segments (multimodal) by accessing OpenStreetMap and GTFS data.

Trillium

How do we validate OSM data?

OSM data can be validated through the use of control data, through intrinsic quality metrics, or through manual self testing.

1. Use control data for comparison
    a. [How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Data (2010)](#)
        i. A comparison with Ordnance Survey (OS) datasets (the national mapping agency for Great Britain).
        ii. "The analysis shows that OSM information can be fairly accurate: on average within about 6 m of the position recorded by the OS, and with approximately 80% overlap of motorway objects between the two datasets."
            1. Data has improved in the last 10 years, likely reducing the error percentage.
    b. [A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany (2010)](#)
        i. A comparison with navigation company data (miles of roadways, miles of pedestrian paths).
        ii. Better data in bigger cities >>"A significant decrease in the data was observed as the distance from the city center increased".
        iii. Data tends to just get better over time.
        iv. More/better OSM data on pedestrian paths, alleys, etc.
2. Ground Truthing
    a. Lyft's 2020 study "[Ground Truth Evaluation of OpenStreetMap Quality in North American Cities](#)"
        i. Utilized public-health-styled cluster sampling with remote sensing to collect ground truth data for analysis and comparison to current OSM data in 30 major US cities.
        ii. OSM has a very high quality road network in the cities sampled.
        iii. The OSM community keeps the map reliably up to date in these areas.
        iv. Gaps in data are usually a result of recent changes in the real world versus historically unmapped features.
        v. OSM onramp signage and lane annotation can be improved in some cities.
        vi. This survey methodology is potentially valuable for future map quality groundtruth studies
3. [Intrinsic quality metrics](#) - For places where authoritative control data is tough to get or nonexistent.
    a. The relative density of contributors and their edits in a given area.
    b. The recency of when the edit occurred.
    c. A "Very different measurement of quality (does not measure completeness or positional accuracy), but provides insight into where/when contributions are occurring, a potential proxy for...how many eyes validated that part of the map."
4. Self Testing
    a. Process

Trillium

     i.   For any given area, determine if on average:
1. OSM has less detail than Google Maps (GMaps).
2. OSM has the same detail as GMaps.
3. OSM has more detail than GMaps.
4. The assets on OSM are accurate, using GMaps as reference.
   a. *We did not find a single OSM asset that was inaccurate when measuring against GMaps.*

    ii.   Pick an area to compare (e.g. Cities within NWOTA's service area, plus a random spot along a US highway).

    iii.   Find a spot within an area that has features to compare, and match lat/lon and zoom level of OSM and GMaps.
1. Zoom level should be adjusted as analysis is conducted so that all possible assets are displayed.
   a. ex: GMaps shows preference with what businesses display depending on the zoom level.

    iv.   Start with roads and, using GMaps as a reference, assess if OSM roads:
1. Exist where GMaps roads do
2. Match GMaps' shapes
3. Have names, and if they match GMaps
4. Exist where GMaps roads do not (Using GMaps' satellite imagery for reference)
5. Make note of percentage of analysis area where OSM quality is worse than GMaps, the same, and exceeds.

    v.   Compare buildings, using the same process in step 3.
1. In cases where an asset exists in OSM, but differs, whether by shape or name, determine which is correct (e.g. a building may have changed businesses recently).

    vi.   Compare other assets (i.e. greenspace, parking lots, restrooms, etc.), using the same process in step 3.

    vii.   Repeat the process for another 4-5 areas.
1. Make conclusions based on resulting averages of percentages.

    viii.   Test walking directions between GMaps and OSM and compare results
1. Where are walking directions specifically defined, and where are walkable areas not utilized?

Quality Comparison

Some locations may be more detailed in Gmaps, while others might be more detailed in OSM. Those outliers notwithstanding, general trends emerge:

1. GMaps includes more building shapes consistently. They likely have an algorithm that

automatically detects buildings from satellite images. This could explain why it draws shapes for sheds in the middle of nowhere. This kind of information is unneeded for common trip planning, and in most cases, not useful.

2. The cases where GMaps includes more building shapes than OSM are usually in residential zones.
3. OSM is more inconsistent with building shapes and names, but where they exist, they are more detailed than GMaps.
   a. Hand-drawn by someone invested in the stewardship of the map, likely a local.
4. Street shapes and street name presence is virtually the same, though OSM includes more private drives, service roads, and other secondary roads.
5. OSM maps more secondary pathways, roads, and walkways. As a result, OTP's walking routes are more direct. The algorithm in both applications will choose pedestrian-only paths when available. Because OSM has more mapped walkways, OTP directions are also safer. While there are some cases of GMaps providing better walking directions than OTP, the vast majority of cases show OTP to be as good, or better.

Cities Tested

1. Keizer, OR
   a. Buildings, roads, and assets **OSM same as GMaps**
      i. With the exception of building shapes (especially non commercial/non municipal), which GMaps more consistently represents.
   b. Walking directions **OTP better than GMaps**

2. Seaside, OR
   a. Buildings, roads, and assets **OSM more detail than GMaps**
      i. With the exception of building shapes (especially non commercial/non municipal), which GMaps more consistently represents.
   b. Walking directions **OTP (slightly) better than GMaps**

3. Banks, OR
   a. Buildings, roads, and assets **OSM more detail than GMaps**
      i. OSM is more detailed on all accounts.
   b. Walking directions **OTP better than GMaps**

4. Yachats, OR
   a. Buildings, roads, and assets **OSM less detail than GMaps**
      i. GMaps has more building names (and shapes, per the norm), but street shapes and names are equal in detail.
   b. Walking directions **OTP (slightly) better than GMaps**

5. Rainier, OR
   a. Buildings, roads, and assets **OSM more detail than GMaps**
      i. OSM is more detailed on all accounts
   b. Walking directions **OTP better than GMaps**

Trillium

6. US-30 and Valley Creek Rd, between Astoria, OR and Clatskanie, OR
   a. Buildings, roads, and assets-**OSM more detail than GMaps**
      i. Except for building shapes
   b. Walking directions **OTP better than GMaps**

Keizer, OR OTP uses an unpaved road north of the park
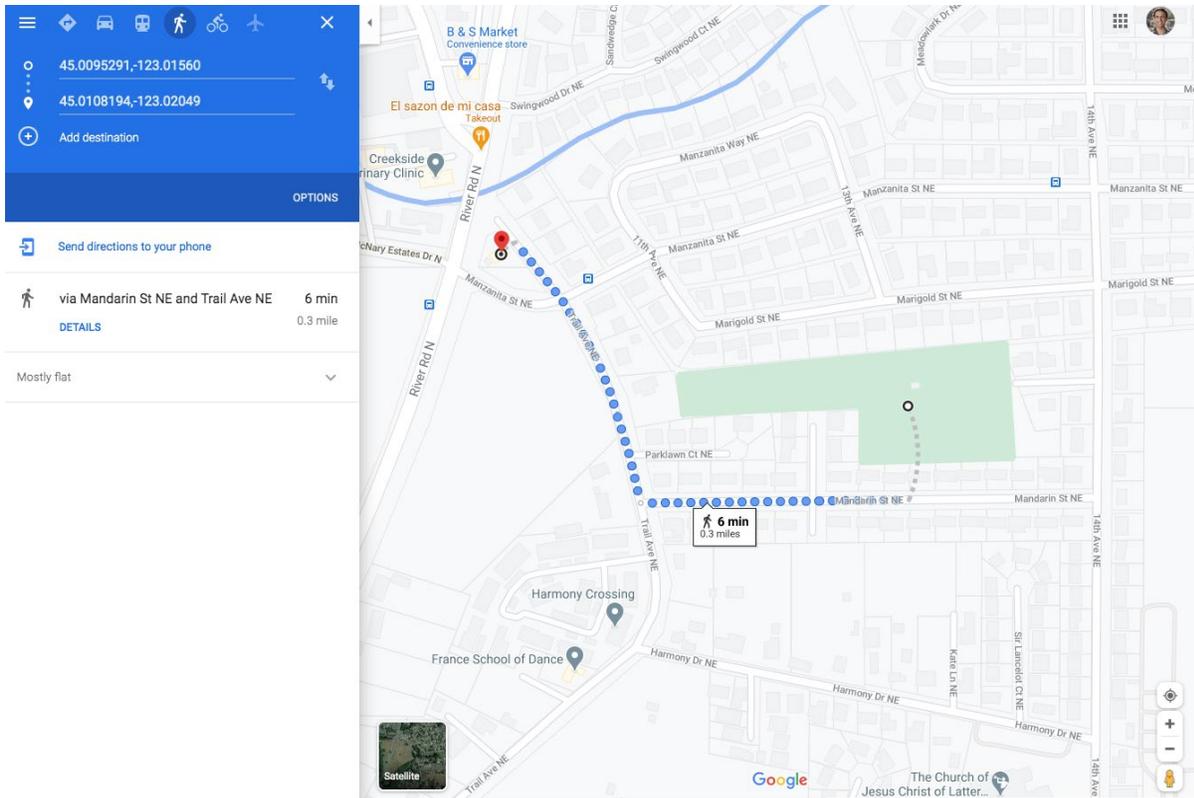
Keizer, OR: GMaps utilizes a longer walk south of the park
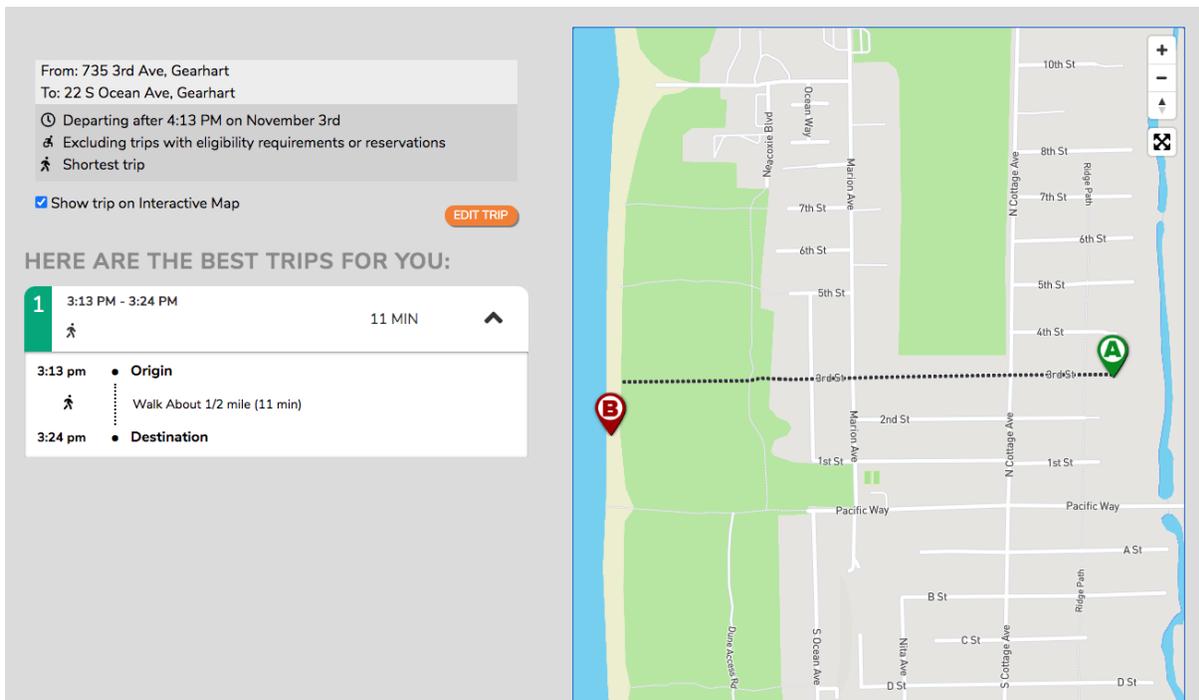
Gearhart, OR: OTP uses a path to the beach off 3rd St.
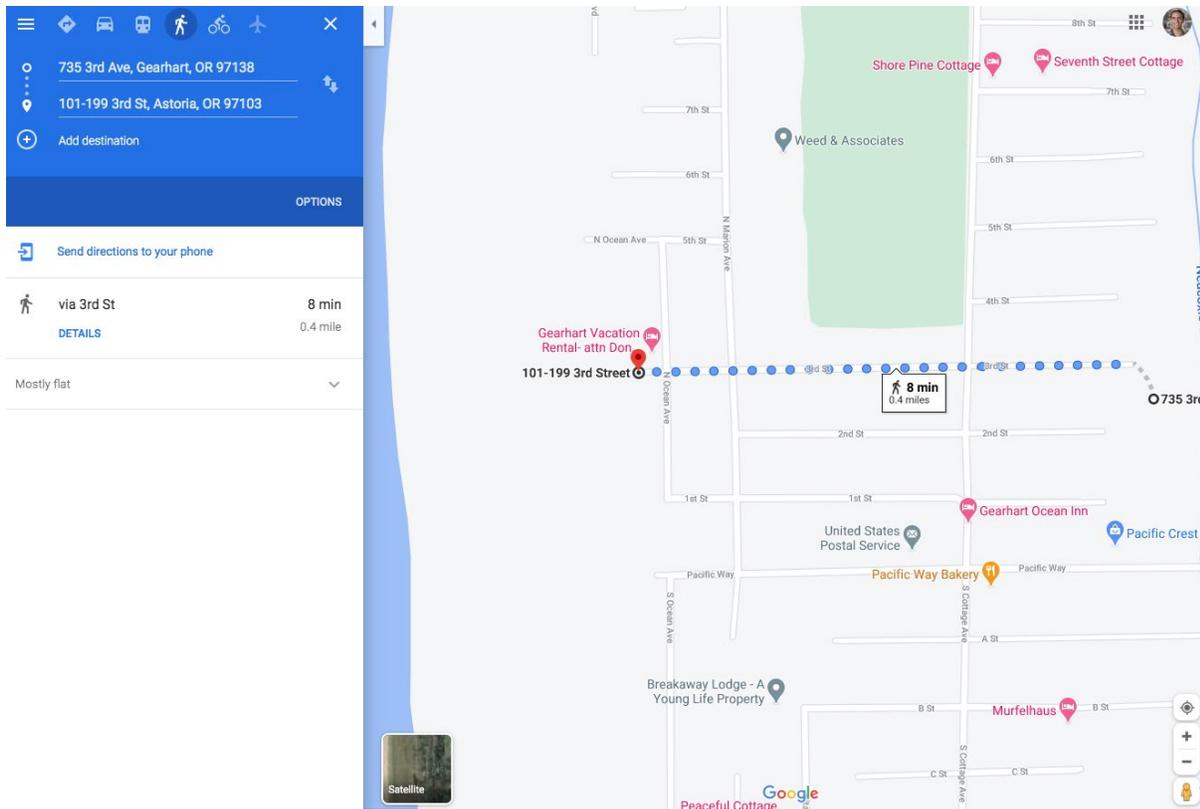
Gearhart, OR: [GMaps](#) does not recognize the path or the beach as accessible. The farthest it plans is the end of 3rd St.

Biking, Walking, and Accessibility

Sidewalks, walkways, curbs, and bike lanes are more difficult to validate than other types of infrastructure. Verification often necessitates ground truthing, as available satellite data is almost never contemporary enough to allow informed vetting. Consequently, these features are rarely as well-documented and maintained as the road network.

1. Sidewalks/walkways
   a. How are they mapped in OSM?
      i. As a feature by tagging the highway (any roadway) with sidewalk=both/left/right/no as appropriate for those sections of sidewalk that are parallel with the carriageway.
         1. Attributes of sidewalks are sometimes added as further tags on the highway, for example:
         2. sidewalk:left:width=3 m
         3. sidewalk:left:kerb=raised
         4. sidewalk:both:surface=paving_stones
      ii. Mapped as a separate way using highway=footway.

    b. Walking directions: How are they created if a road isn't tagged with sidewalk info, or if a separate sidewalk is not created as a separate way?

        i. OSM assumes, per most state laws including OR[1] and WA[2], that all roads are legally accessible to pedestrians except in certain cases where it is assumed that pedestrians are not allowed, like highways/freeways, unless otherwise stated.

    c. How does mapping method affect accessibility?

        i. Curbs (or kerbs in the UK-developed OSM) at crossings are assigned as a tag to a sidewalk.

            1. A sidewalk must already be created as a separate way, or you must create it first.

            2. No separate sidewalks=no identification of the type of curb that exists on each street corner or crossing

    d. Coverage in OSM

        i. Because of the high level of detail, time consuming data entry, and limited availability of up-to-date imagery, sidewalks and curbs are not extensively mapped in OSM or GMaps even in metropolitan areas.

2. Biking and bike lanes

    a. Similar to sidewalks and curbs, bike lanes are difficult to identify in a timely manner. The responsibility falls almost solely on locals to map, save for older features visible in satellite imagery. See CycleOSM in resources below.

---

[1] T. (Ed.). (2016, November 09). Pedestrian Rights On Roadways. https://www.tcnf.legal/pedestrianrights/
[2] Washington State pedestrian laws. (2019, December 10). https://wsdot.wa.gov/travel/commute-choices/walk/pedestrian-laws
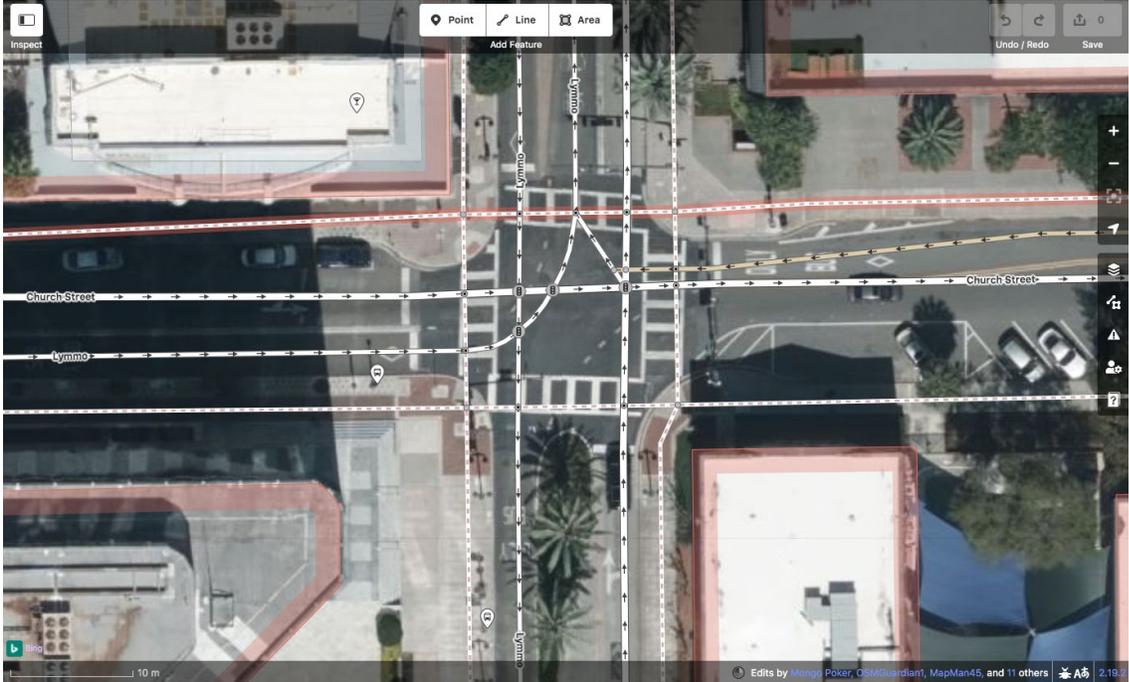
Sidewalks mapped in Orlando, FL (dashed lines)

A few footpaths marked near the beach in Cannon Beach, OR, but otherwise no sidewalks are separately mapped.

Validation Resources

1. [Is OSM up-to-date?](#)
   a. OpenStreetMap history for intrinsic quality assessment.
   b. Shows how recent nodes and ways were edited.
   c. Move map then click the "fetch data" button in upper right to get updated info. The box in the upper right corner shows the oldest edit in view.
   d. Is [OSM up-to-date? Github](#)
2. [Osmose Map](#)
   a. A quality assurance tool that detects issues in OpenStreetMap data.
   b. Most errors are minor.
3. [Latest OpenStreetMap Edits per Tile](#)
   a. Displays edits within the last 2 weeks
4. [Map of users](#)
   a. Displays the level (beginner to expert) of the mapper who made each edit in the last 6 months
5. [OSM wiki validation tools](#)
   a. A list of several more open source tools.

6. [OSM History](#)
   a. Displays a geotagged changelog for the map extent.
   b. View edits that date back as far as you want (though you have to manually load additional logs in batches).
7. [CycleOSM](#)
   a. Bicycle-oriented OSM map highlighting biking lanes, roads etc. tagged in OSM.
   b. Allows for easy viewing and identification of areas that have, or are missing the appropriate bike tags.
   c. You can see scant tagging of bike lanes etc. within the NWOTA service area. This may be because there aren't many bike friendly streets or designated lanes.
8. Contacting mappers
   a. The OSM ethos is about community, and fostering relationships towards the goal of better data. Accordingly, it is easy to connect with other users.
      i. Each user's edits are logged and tagged with metadata, and you can link directly to the user profile where you can contact them and even add them as a friend. See the example below.

OpenStreetMap ☰

amarajz

Edits 7,576 | Map Notes | Traces 0 | Send Message | Diary 0 | Comments | Add Friend | Report this User

Mapper since: July 03, 2019

I work for Amazon Logistics. At Amazon Logistics, we've been utilizing OSM in some cases related to our delivery programs. In connection with those delivery programs, we have collected information that we think is valuable to the OSM community such as names and info about new roads that are not currently in the map today, new data on turn restrictions, and road connectivity, to name a few. When we hear feedback, we've been editing to provide that information for the benefit of the entire OSM user community. If you have more questions, please contact osm-edit-escalations@amazon.com

Mapping Resources

1. OpenStreetCam
    a. OpenStreetCam is a free and open source platform for street level imagery. Anyone can contribute images with a smartphone.
    b. Detects features from the uploaded images such as signs, lanes and road curvature to improve OpenStreetMap.
2. Mapillary
    a. Allows users to capture imagery with any kind of camera. Their software automatically detects a number of street features such as crosswalks, lane markings, and bike lanes.
    b. Open source when used for OSM
        i. https://www.mapillary.com/osm

Who's editing In NW Oregon?

1. There isn't a fully-fledged, coordinated effort.
    a. Some are hobbyists who make edits sporadically, even some who edit at a national or international level with no real pattern or schedule to editing.
2. Amazon Logistics personnel are making substantial edits.
    a. Working at a national level to aid in deliveries
3. Common editors in NWOTA service area
    a. Utible
        i. Vast majority of edits in the last few months. Cannon Beach and northward
    b. Sushredd
    c. StanB
        i. Pacific City to Lincoln City area mostly
4. Edit Areas
    a. Not surprisingly, the frequency/density of edits appears highest in the northern extent of the NWOTA service area.

Trillium

    b. Astoria is edited much more (e.g. ~19 changesets by 4 people in a 2 week period) than the Newport area (0 in the same period).

Who uses OSM maps?

1. Locally
   a. Salem Cherriots
      i. Cherriots built a new bus transit center in 2013, and the land was re-developed to the extent that it did not show up on the map correctly.
      ii. Cherriots' staff used OSM editing tools to draw the roads and walking ways for the transit center.
      iii. Upon the next refresh of OSM tiles, the new roads were visible in their OpenTripPlanner (powered by OSM).
         1. Roads/ways were available for driving and walking directions.
      iv. GMaps was not updated until much later.
   b. TriMet
      i. TriMet has an extensive OSM editing program, which it uses to support the OpenTripPlanner software it relies on for customer trip planning [more information].
2. Globally
   a. Amazon
      i. Amazon uses OpenStreetMap for deliveries. They collect information valuable to the OSM community such as roads not currently mapped, turn restrictions, and road connectivity, to name a few.
   b. Lyft
      i. Lyft uses OpenStreetMap as a base for all its mapping software. See results of their 2020 Lyft's 2020 study "Ground Truth Evaluation of OpenStreetMap Quality in North American Cities" referenced in the validation section.
   c. Apple
      i. Apple volunteers contribute to the Missing Maps project and other data improvement projects in OSM around the world.
   d. Facebook
      i. Facebook has used OSM for its maps since 2018, and is a gold member of the OSM Foundation. Their AI-assisted road tracing and RapiD editor allows OSM users to map unidentified roads around the world.
   e. Microsoft

      i. Microsoft was the first company to provide OSM with [aerial imagery](#) and in 2018 they uploaded [124 million building footprints](#) in the US. They have their own [team](#) working to improve data in OSM.

f. Esri

      i. [Live OpenStreetMap data in ArcGIS](#)

        1. Esri now generates vector tiles that are updated every 3 weeks with the latest OSM data.

g. Baidu Maps

h. Foursquare

i. Pinterest

j. Pokemon Go

k. Snapchat

l. Air France

m. Alaska Airlines

n. Deutsche Bahn

o. Grab

p. SNCF (French rail agency)

q. Uber

r. CARTO

s. Digital Globe

t. Garmin

u. Mapbox

v. Telenav

w. BBC

x. Financial Times

y. The Guardian

z. National Geographic

aa. New York Times

bb. US News & World Report

cc. Wall Street Journal

dd. Washington Post

ee. And many more governments, countries, universities, and foundations

Reflections and recommendations

Trillium began this review with the expectation that OSM data would not be accurate enough for ODOT's purposes, and that we would need to produce substantial data for OSM to make trip planning better. Our review has found the opposite. OSM data has proven to be reliable. There is a global editing system in place, and it appears to be working. To aid in this effort locally, ODOT should invest resources to contribute data in targeted areas to enhance trip planning and accessibility. The following recommendations outline how ODOT should approach OSM data going forward.

- Trust OSM, but also conduct periodic spot validations and be prepared to make any necessary edits.
- OSM data should be considered for other purposes besides OTP.
  - Examine what it would look like to adopt OSM as an official map product, especially given Esri users' new ability to work with live [OSM data in ArcGIS](). Transitioning to OSM could reduce time and money spent updating internal datasets by instead leveraging the comprehensive, contemporary, and open source information provided by OSM.
- Improve and add to OSM data in ways that will be most beneficial to all riders, starting with the NWOTA service area.
  - Update walkways where a path/sidewalk exists on public property, focusing on areas other than adjacent to a road.
    - [Transit Pedestrian Connectivity]() by Evan Siroky of IBI
  - Update road tags in places where people should definitely not walk and where walking is prohibited to discourage routing through these areas.
  - Consider an informal budget assessment by interviewing relevant staff at TriMet
  - Update the road network with bike lane information where applicable.
    - Explore the capabilities and efficacy of OpenStreetCamera and Mapillary in capturing information on the ground and automating the update process.

Trillium