



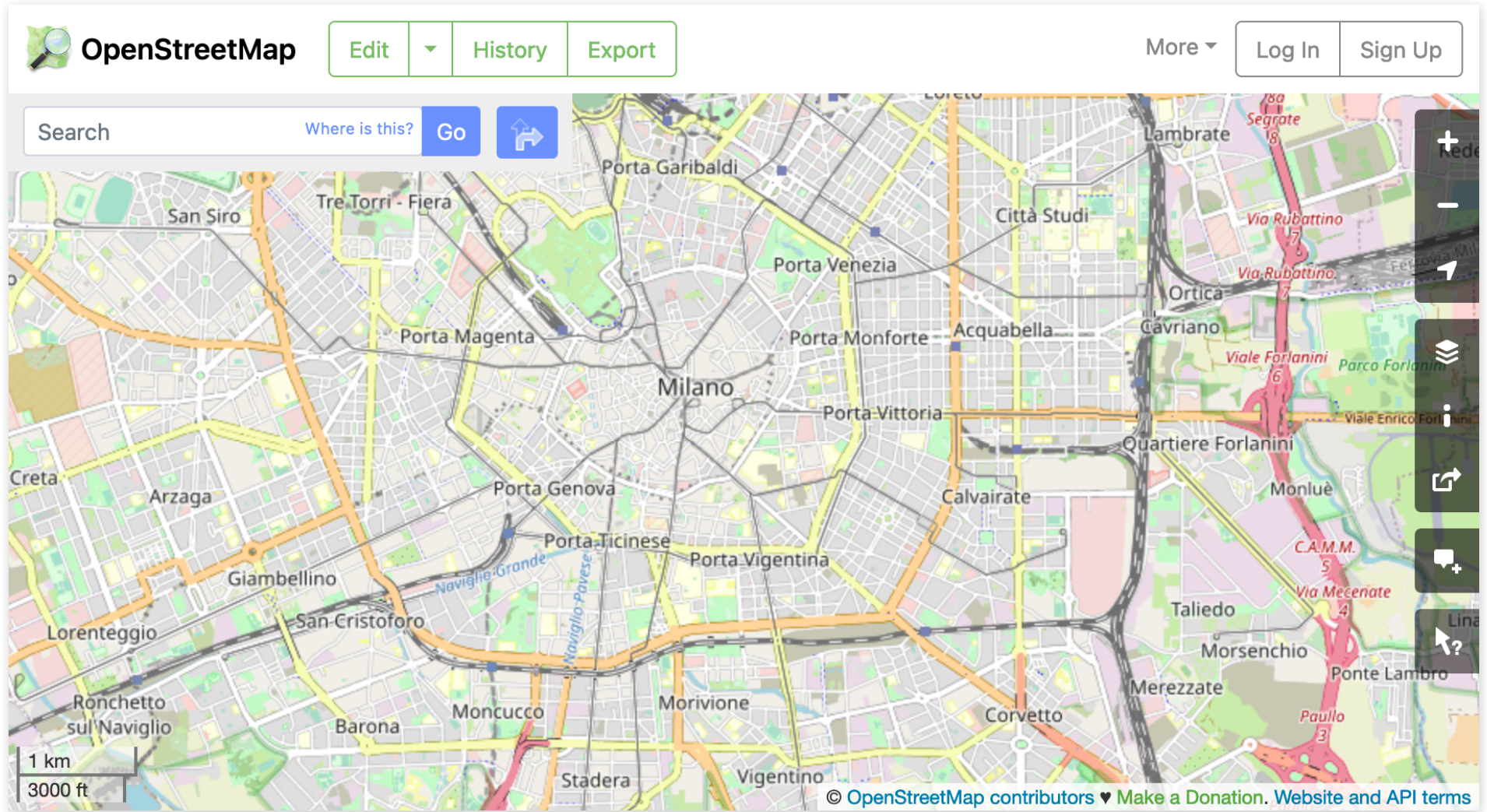
POLITECNICO
MILANO 1863

Towards the evaluation of OpenStreetMap intrinsic temporal accuracy and up-to-dateness using historical data

Author: Francesco Frassinelli
Professor: Maria Antonia Brovelli
Co-advisor: PhD Marco Minghini

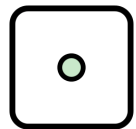
22 July 2022

OPEN STREET MAP (OSM)



OPEN STREET MAP (OSM) - DATA MODEL

OSM element



Node



Way (open)

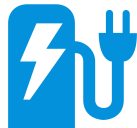


Way (closed)



Relation

Real-world example



Charging facility
for electric vehicles



Railway



Playground



Turn restriction

Tags example

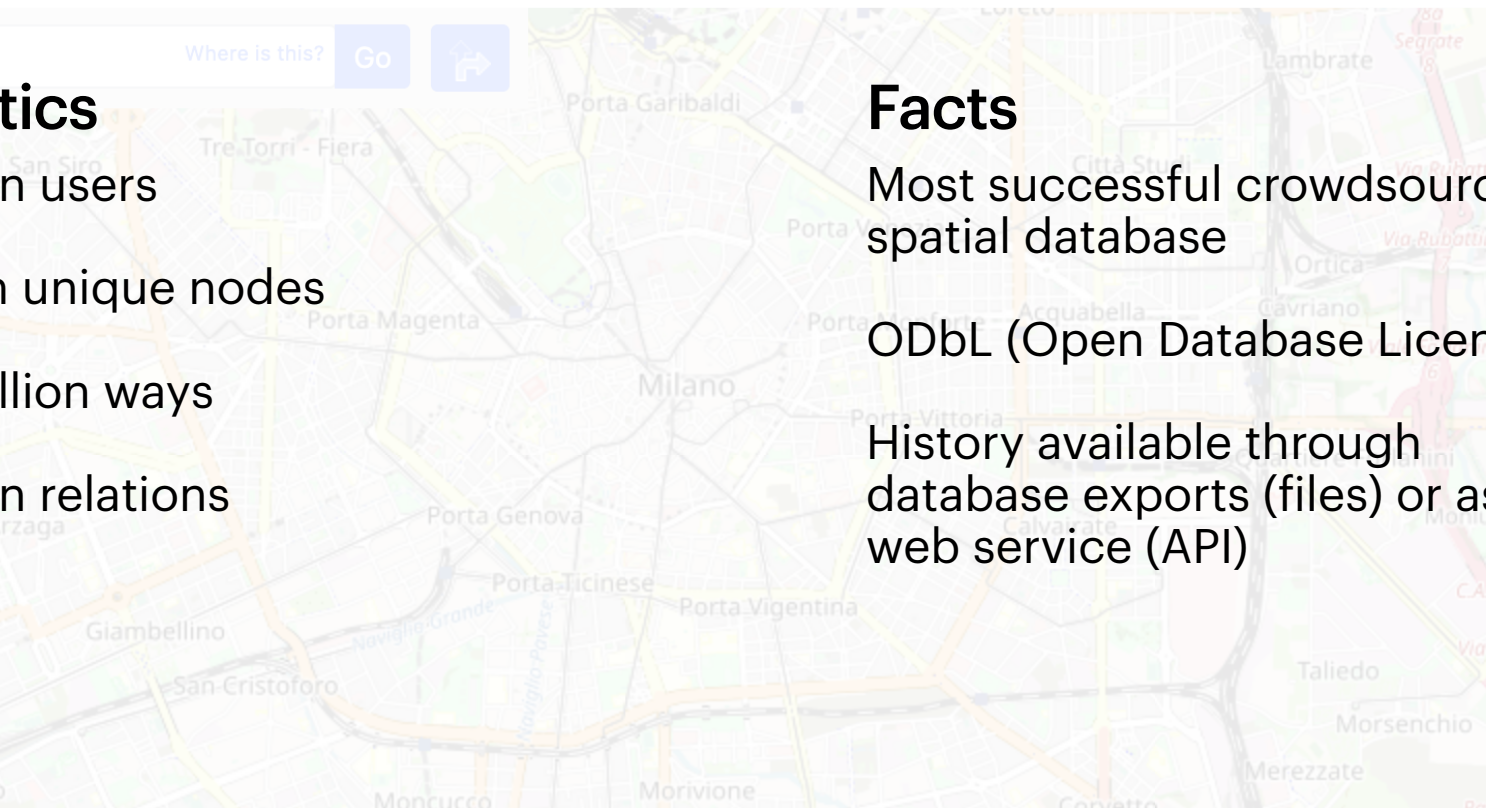
amenity=charging_station
socket:type2=1

route=railway

leisure=playground
wheelchair=yes

from=...
to=...
restriction=no_right_turn

OPEN STREET MAP (OSM)



Statistics

- 8 million users
- 7 billion unique nodes
- 800 million ways
- 8 million relations

Facts

- Most successful crowdsourced spatial database
- ODbL (Open Database License)
- History available through database exports (files) or as a web service (API)

PROBLEM BACKGROUND

01

Estimate the quality of OpenStreetMap

in order to overcome barriers that slow down its adoption.

02

Different possible criteria

spatial accuracy, lineage, semantic consistency, temporal completeness.

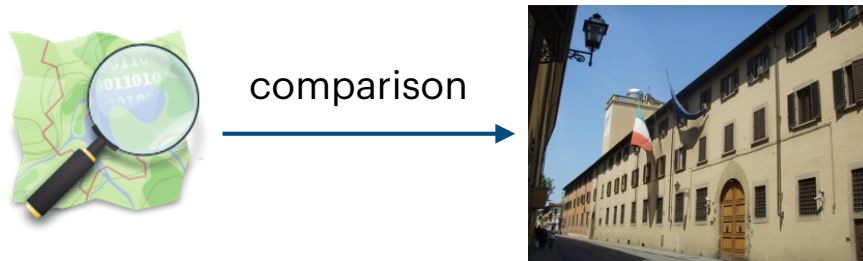


PROBLEM BACKGROUND

Extrinsic

Compare OSM to an external database.

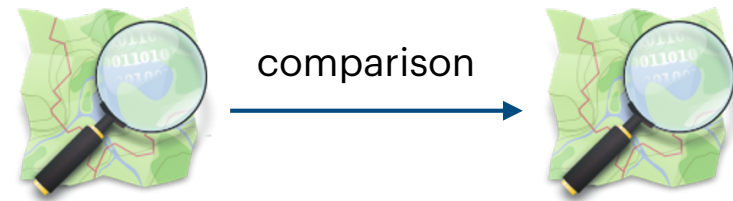
Example: official map



Intrinsic

Compare OSM to itself, to reality or to OSM wiki.

Example: tag violation



GOALS

Evaluate

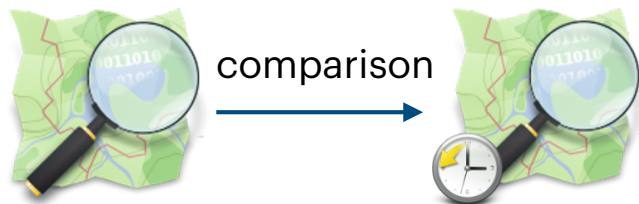
- the feasibility of estimating the intrinsic temporal accuracy (and doing it with limited resources);
- the existence of already available user-friendly solutions.

Develop

a new solution if needed,
or improve existing ones.

Suggest

new and different technical approaches, to further improve how the data is gathered and processed.



SHARING PRELIMINARY RESULTS ON FEASIBILITY - CONFERENCES (2018)



FOSS4G-IT

21 February 2018
Rome



Merge-IT
24 March 2018
Turin

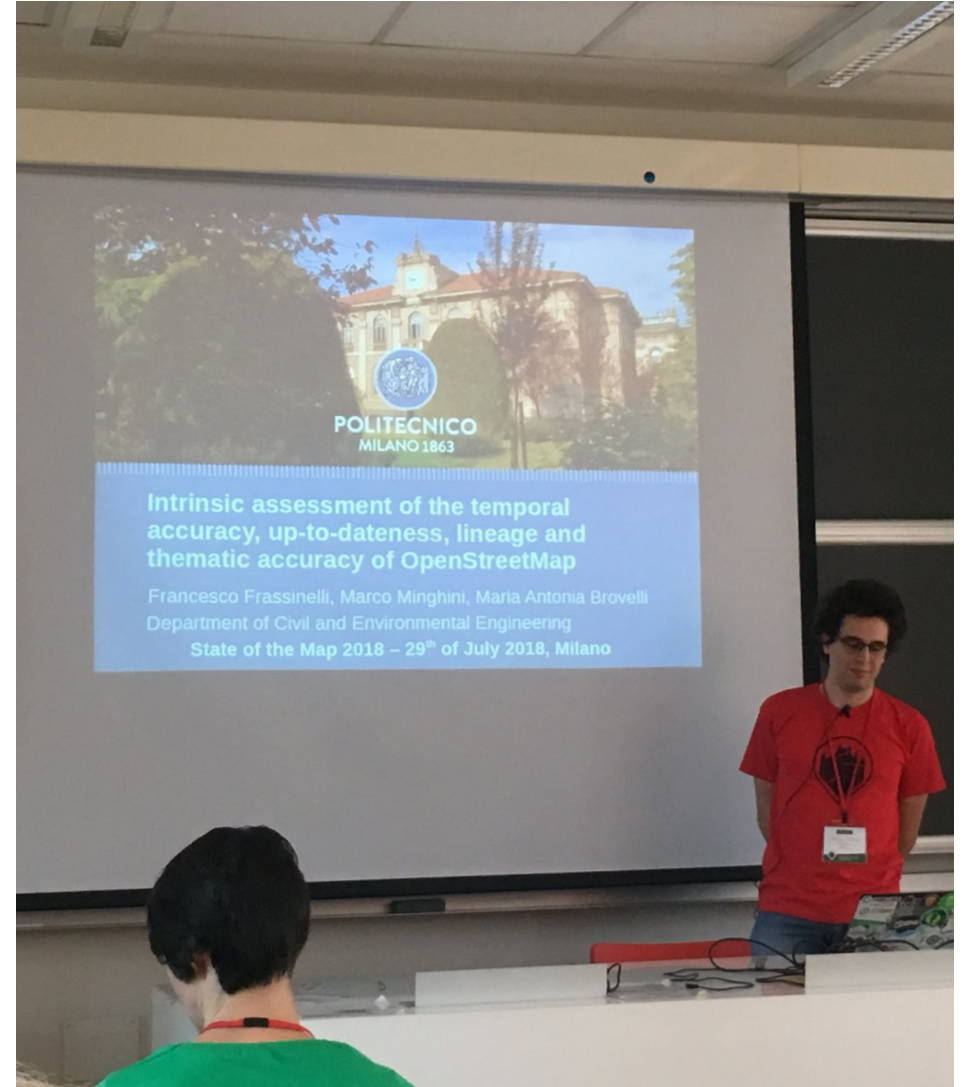


State of the Map

29 July 2018
Milan



FOSS4G
31 August 2018
Dar es Salaam



CHALLENGE: HOW TO HANDLE AND PROCESS THE DATA?

	Solution A: Custom script	Solution B: Web application
Data source	OSM planet history file	OSM API
Automatic	Partially automated	Fully automated
User friendly	No	Yes
Storage needed	>100 GB	<1 GB
Area	100.000 km ²	1 km ²
Processing time	Few hours	Few seconds
Data aggregation	Yes	No

Contributors
density.
FOSS4G-IT
2018



Average
creation time.
State of the map
2018



INSIGHTS: TRENDS AND OUTLIERS

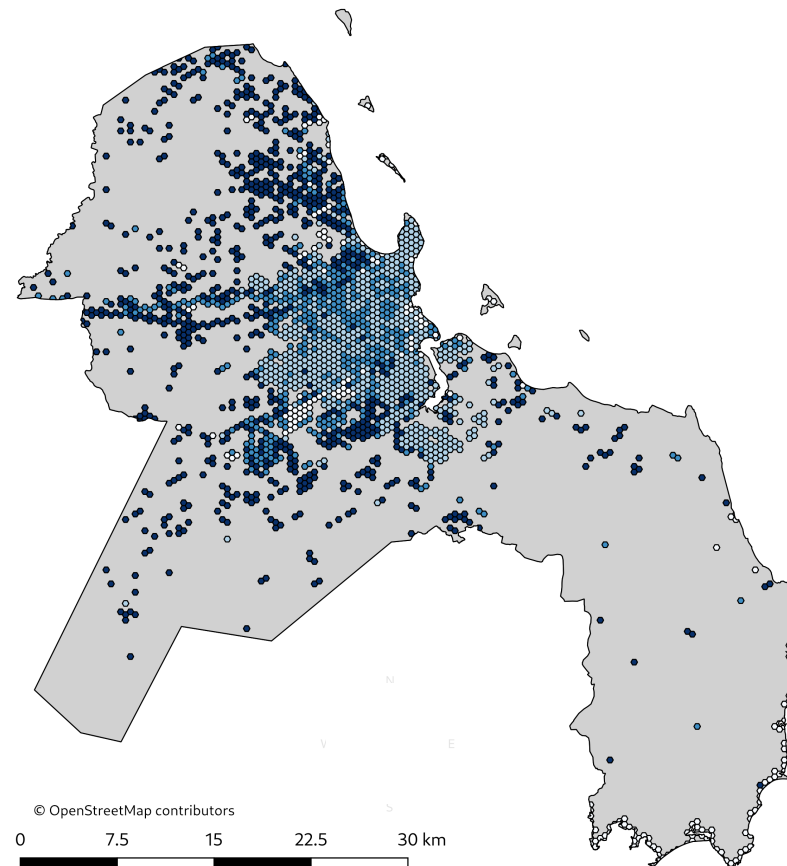
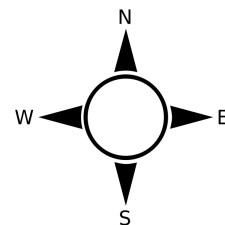
Average
last edit.
FOSS4G
2018



FOSS4G Conference paper published by
*ISPRS - International Archives of the
Photogrammetry, Remote Sensing and
Spatial Information Sciences*

Authors: M. Minghini,
M. A. Brovelli, F. Frassinelli

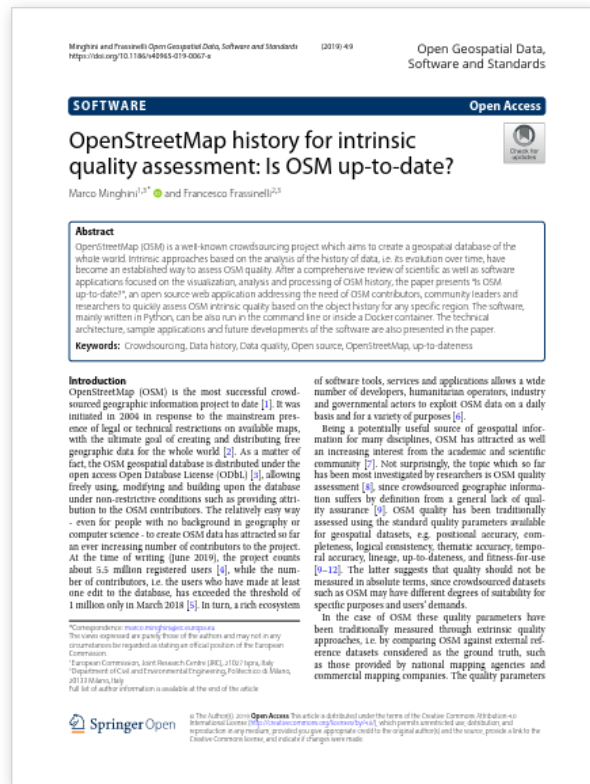
Last edit (average)
 1 Jan 2007 $\leq x \leq$ 31 Dec 2014
 1 Jan 2015 $\leq x \leq$ 31 Dec 2015
 1 Jan 2016 $\leq x \leq$ 31 Dec 2017
 1 Jan 2018 $\leq x$
 Administrative boundary



<https://is-osm-uptodate.frafra.eu/>



STATE OF THE ART & WEBAPP PROTOTYPE - PUBLICATIONS (2019)



Peer-reviewed paper published by
Open Geospatial Data, Software and Standards - Springer Open

Authors: M. Minghini, F. Frassinelli

Table 1 Classification of the available software applications based on OSM history according to the overall aim, type, history source and license

From: *OpenStreetMap history for intrinsic quality assessment: Is OSM up-to-date?*

Aim	Name	Reference	Type	History source	License
Visualization	openstreetmap.org	[64, 123]	Web app	OSM API	GPL 2
	Achavi	[65, 124]	Web app	Overpass API	MIT
	OSM Deep History	[23, 125]	Web app	OSM API	BSD
	OSM History Explorer	[68]	Web app	OSM Planet	Closed
	OSM History Renderer*	[75]	Application	OSM planet	BSD 2
	OSM History Viewer (osmmhv)	[22, 126]	Web app	OSM API	AGPL 3 + MIT/PLD
	OSM History Viewer (pWu version)	[26, 127]	Web app	OSM planet	Apache 2
	OSMCha	[70, 71]	Web app	OSM planet	GPL 3
	OSMapi*	[66, 87, 128]	Web app	OSM planet	MIT
	OSMvis	[29, 80, 129]	Web app	OSM planet	GPL-3
	Show me the way	[28, 130]	Web app	Overpass API	MIT
	Visualize Change	[76, 131]	Web app	OSM planet	BSD 3 clause
	Who did it?	[72, 132]	Web app	OSM Planet	WTFPL
Statistics	OSM Latest Changes	[81, 133]	Web app	OSM API	Public domain
	Is OSM up-to-date?	[100, 102]	Web app	OSM API	AGPL 3
	OSM wiki - Stats	[82]	Web page	OSM planet	Closed
	OSMStats	[83]	Web app	OSM planet	Closed
	How did you contribute to OpenStreetMap?	[84]	Web app	OSM planet	Closed
	osmstats.stevecoast.com	[87, 134, 135]	Web app	OSM planet	Source available
	OSM Tag History	[88, 136]	Web app	OSM planet	ISC
	osmism	[85, 89]	Web app	OSM planet	Closed
	Brave Mappers	[89, 137]	Web app	OSM planet	MIT
	iOSMAnalyzer*	[52, 138]	Application	OSM planet	GPL 3
Visualization and statistics	OSM Analytics	[90, 139]	Web app	OSM planet	BSD 3 clause
	OSM Live Changes	[91, 140]	Web app	OSM planet	Apache 2
	Oshome Dashboard	[92]	Web app	OSM Planet	Closed
	EPIC-OSM*	[95, 141]	Framework	OSM planet	MIT
	OSM Data Classification	[99, 93, 142]	Framework	OSM planet	GPL 3
Analysis	OSM Data Classification	[99]	Tool	OSM planet	MIT
	OSMesa	[94, 143]	Framework	OSM planet	Apache 2
	OSM History Importer*	[73]	Tool	OSM planet	BSD 2 clause
Conversion	postgres_osm_pbf_fdw	[96]	Tool	OSM planet	PostgreSQL
	OSM Parquetizer	[97]	Tool	OSM planet	Apache 2
	Osmium	[98, 144]	Tool	OSM planet	GPL 3

* Deprecated

Francesco Frassinelli
Politecnico di Milano
Verified email at mail.polimi.it - [Homepage](#)

ARTICLES CITED BY PUBLIC ACCESS

SORT ▼ CITED BY

OpenStreetMap history for intrinsic quality assessment: Is OSM up-to-date? 35
M Minghini, F Frassinelli
Open Geospatial Data, Software and Standards 4 (1), 1-17, 2019

Number of citations: 35



CHALLENGE: HOW TO HANDLE AND PROCESS THE DATA?

	Solution A: Custom script	Solution B: Web application	Best from both?
Data source	OSM planet history file	OSM API	
Automatic	Partially automated	Fully automated	
User friendly	No	Yes	
Storage needed	>100 GB	<1 GB	
Area	100.000 km²	1 km2	
Processing time	Few hours	Few seconds	
Data aggregation	Yes	No	



CHALLENGE: HOW TO HANDLE AND PROCESS THE DATA?

300 commits later...

Performances

Tiled data caching (backed by Redis), data pipelining, compression, Ohsome API, parallelization, cloud autoscaling, various optimizations...

+

New features and improvements

Clustering, automated testing, statistics, QGIS integration, higher reproducibility...

LATEST VERSION - v2.2 (June 2022)

Search or jump to... Pull requests Issues Marketplace Explore

frafra / is-osm-uptodate Public

Unpin Unwatch 2 Fork 3 Star 17

<> Code Issues 3 Pull requests 11 Actions Wiki Security Insights Settings

master 15 branches 13 tags Go to file Add file Code

frafra Update README.md ✓ 482c225 15 days ago ⌚ 379 commits

.circleci	Reorganize structure and documentation	2 months ago
.github	Build Redis image	2 months ago
docs	Add support for GeoJSON boundaries	last month
redis	Enable Redis persistence on fly.io	2 months ago
server	Add missing quote	23 days ago
tests	Minor Docker improvements	last month
web	Improve settings/filters UI	last month
.gitignore	Use Webpack	9 months ago
.pre-commit-config.yaml	Update .pre-commit-config.yaml	2 months ago
Dockerfile	Improve Dockerfile	last month
LICENSE	Create LICENSE	5 years ago
README.md	Update README.md	15 days ago
docker-compose.yml	Minor Docker improvements	last month
flv.toml	Minor Docker improvements	last month

About

Find outdated nodes in OpenStreetMap

is-osm-uptodate.frafra.eu

react python redis openstreetmap osm leaflet python3 seleniumbase ohsome

Readme AGPL-3.0 license 17 stars 2 watching 3 forks

Releases 13

v2.2.0 Custom boundaries Latest on 5 Jun + 12 releases

Packages 1



CHALLENGE: HOW TO HANDLE AND PROCESS THE DATA?

	Solution A: Custom script	Solution B: Web application	Solution C: Fast web
Data source	OSM planet history file	OSM API	Ohsome API
Automatic	Partially automated	Fully automated	Fully automated
User friendly	No	Yes	Yes
Storage needed	>100 GB	<1 GB	<1 GB
Area	100.000 km ²	1 km ²	100.000 km ²
Processing time	Few hours	Few seconds	Few seconds
Data aggregation	Yes	No	Yes

LATEST VERSION - v2.2 (June 2022) - CLUSTERING

Filters

Filter nodes using [Ohsome selectors](#)

Set

Set custom GeoJSON boundaries

Sfogliala...

Nessun file selezionato.

Criteria

First edit

Last edit

Revisions

Update frequency

Each group of nodes is represented by the value of one of its nodes, chosen

2012-11-13 2022-06-20
Background colour

Edificio 16B, Edificio 14B, Edificio 14A, Edificio 18, Edificio 11 - Architettura, Politecnico di Milano Labora, Edificio 14 - Nave, Via Edoardo Bonardi, Piazza Leonardo Da Vinci Politecnico, Banca Popolare di Sondrio, SeRist, Edificio 4 - Giulio Archi, Edificio 8 - Alessandro Amario

Is OSM up-to-date? | [ohsome API](#) | © OpenStreetMap contributors

LATEST VERSION - v2.2 (June 2022) - NODE DETAILS

Filters

Filter nodes using [Ohsome selectors](#)

example: amenity=*

Set

Set custom GeoJSON boundaries

Sfoggia...

Nessun file selezionato.

Criteria

First edit

Last edit

Revisions

Update frequency

Each group of nodes is represented by the value of one of its nodes, chosen

+

-

Q

2012-11-13

2022-06-20

Background colour

First edit: 2014-03-01

Last edit: 2019-11-25

Revisions: 2 revisions

Update frequency: 0.24 times/year

Attributes:

- [amenity: bicycle_parking](#)
- [bicycle_parking: stands](#)
- [capacity: 40](#)
- [covered: no](#)
- [fixme: capacity: valore approssimativo](#)

[Edit](#) | [History](#) | [Details](#)

Is OSM up-to-date?

ohsome API

© OpenStreetMap contributors

LATEST VERSION - v 2.2 (June 2022) - TILES

Filters

Filter nodes using [Ohsome selectors](#)

example: amenity=*

Set

Set custom GeoJSON boundaries

Sfogliala...

Nessun file selezionato.

Criteria

First edit

Last edit

Revisions

Update frequency

Each group of nodes is represented by the value of one of its nodes, chosen

LATEST VERSION - v2.2 (June 2022) - CRITERIA

Criteria

⏮ First edit

⏭ Last edit

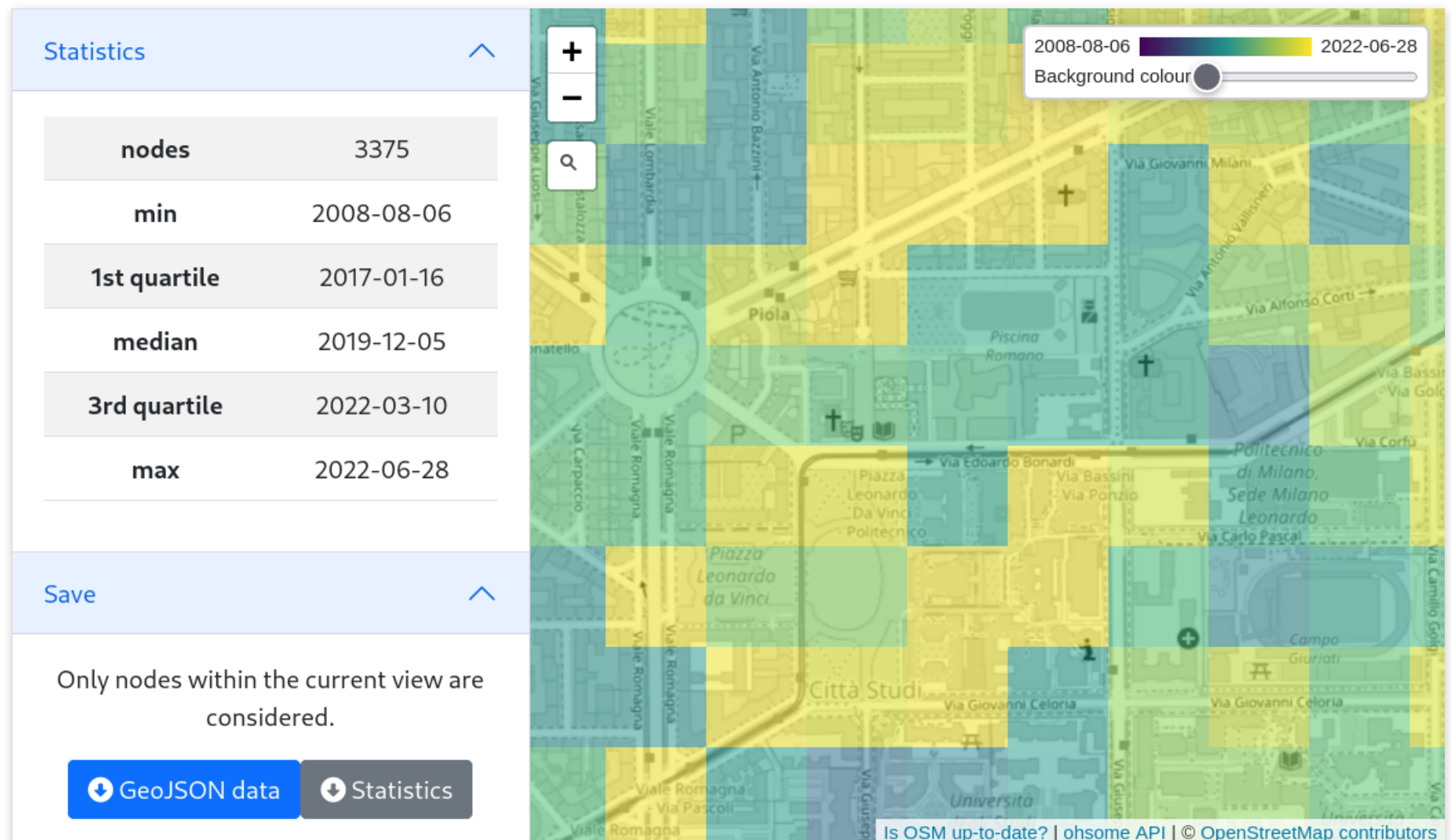
📄 Revisions

🕒 Update frequency

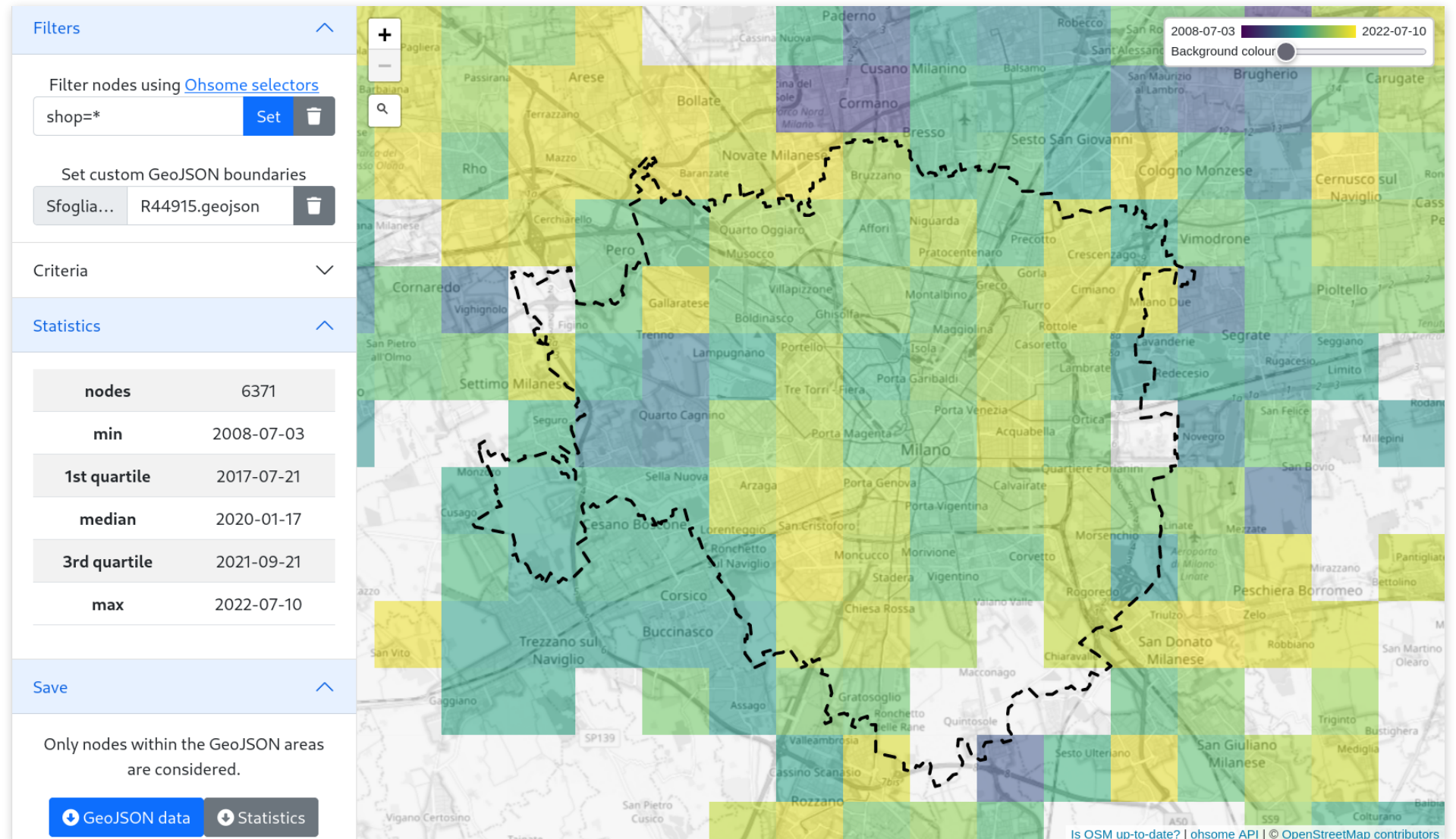
Each group of nodes is represented by the value of one of its nodes, chosen using the following criteria:

Least recently updated
1st quartile
Median
3rd quartile
Most recently updated

LATEST VERSION - v2.2 (June 2022) - STATISTICS AND DOWNLOADS

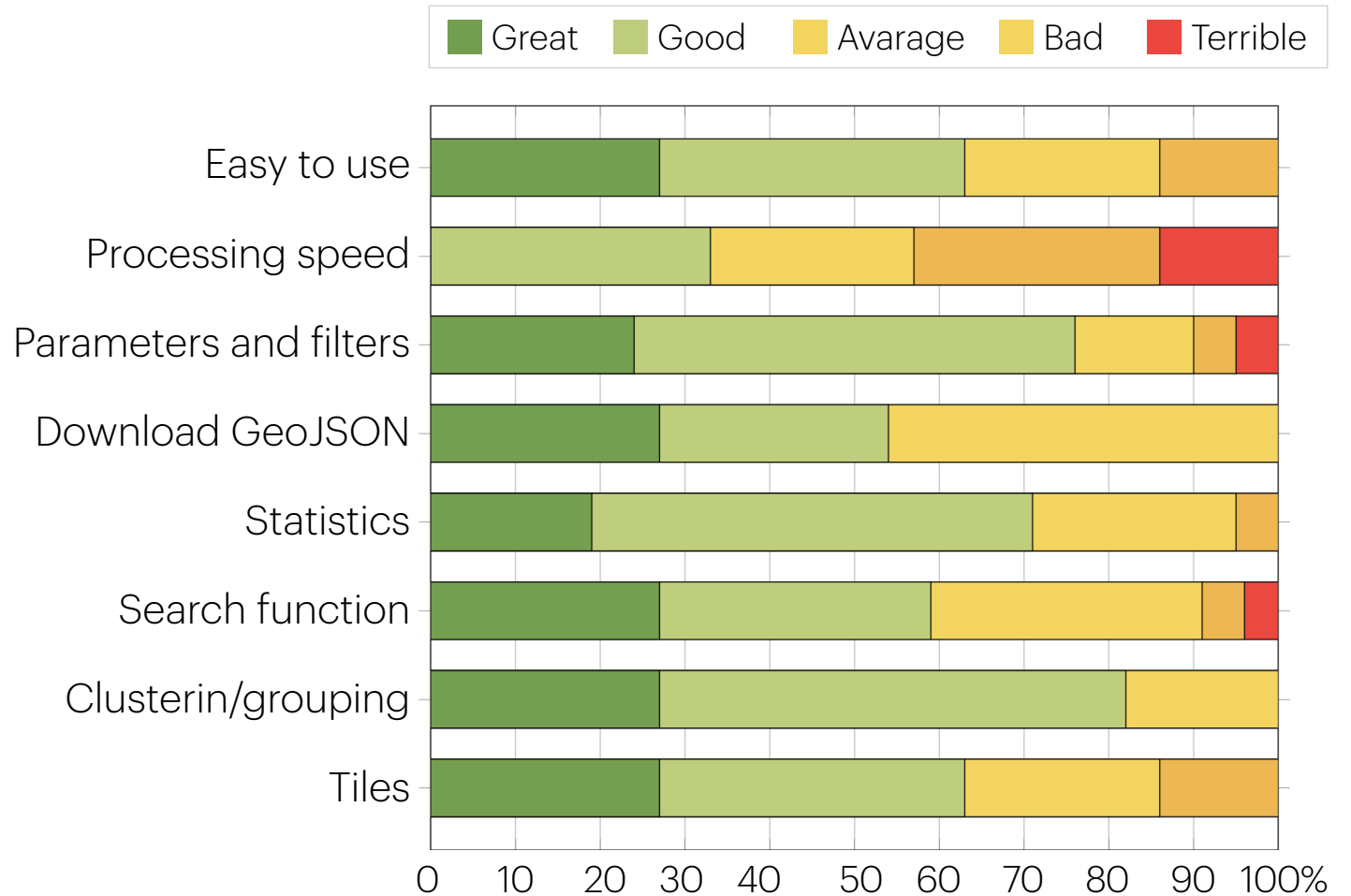


LATEST VERSION - v2.2 (June 2022) - FILTERS AND BOUNDARIES



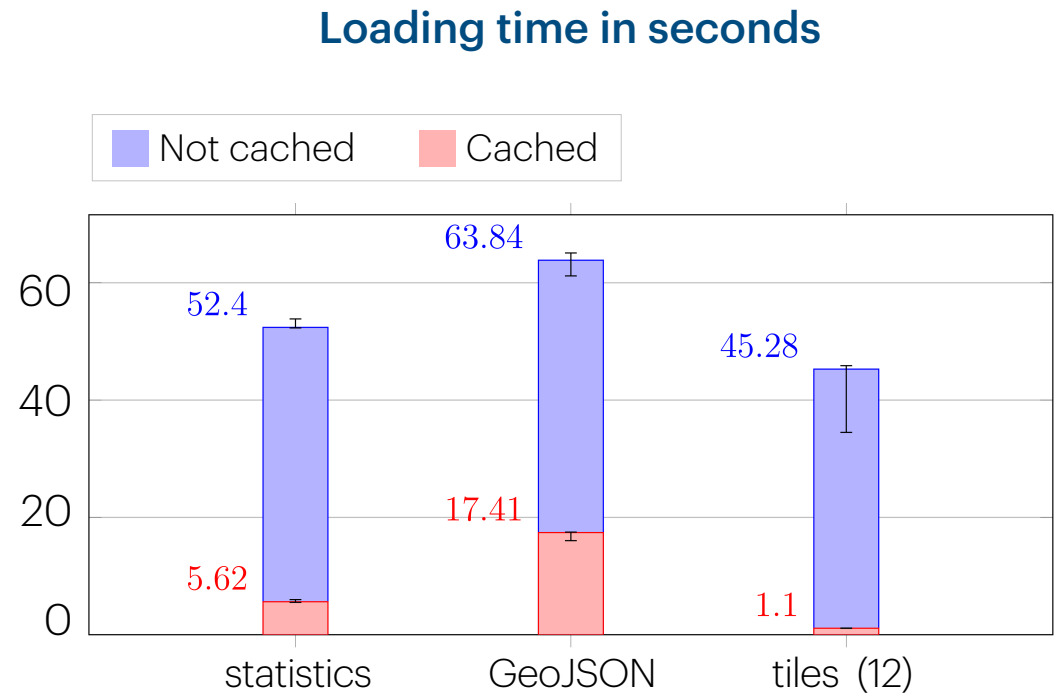
SURVEY

Survey feedback
results from
23 people
(OSM contributors
mostly, as well as
5 with an
academic/research
background)



PERFORMANCE

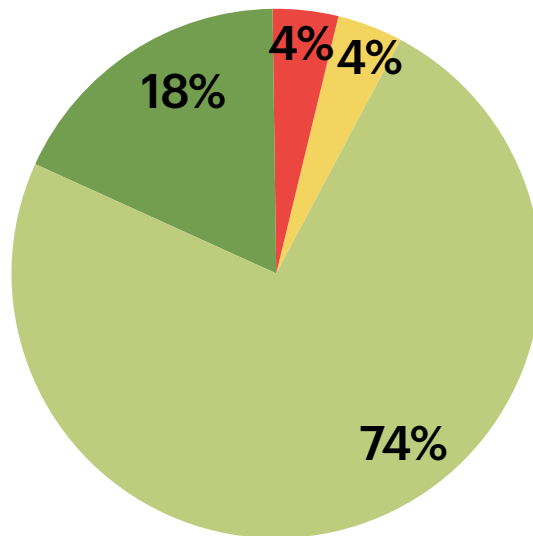
Loading times for the city of Milan.



SURVEY

01

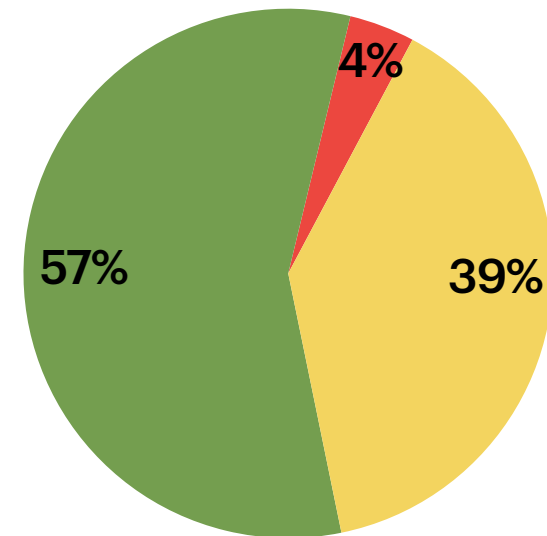
Is it useful?



Very useful Quite useful
Not so useful I am not interested

02

Would you use it again?



Yes Maybe No



FUTURE WORK

Research topics

- Overall quality index

Features

- Improve usability
- Custom boundary for tiles
- GeoTIFF export
- Better error handling
- Better GIS integration

Performances

- Approximate statistics
- Bypass Ohsome (index planet files combined with SQLite Virtual Table)
- Faster rendering
- Faster computation

Hosting

- Higher availability



ANNOUNCEMENT

Is OSM up-to-date? became the **first constant external service using Ohsome**, with almost **100 monthly users**.

A brief interview will be published soon on the blog of the Heidelberg Institute for Geoinformation Technology.



```

29 def _iterate_tiles(multipolygon, start, end, *filters, **headers):
30     if not multipolygon.is_empty:
31         bbox = mercantile.Bbox(*multipolygon.bounds)
32         for tile in bbox_tiles(bbox, Z_TARGET):
33             tile_box = shapely.geometry.box(*mercantile.bounds(*tile))
34             sliced = multipolygon.intersection(tile_box)
35             if sliced.is_empty:
36                 continue
37             fast_comparison = sliced.bounds == tile_box.bounds
38             quadkey = mercantile.quadkey(tile)
39             tiled_data = peek(
40                 get_tile_data(quadkey, start, end, *filters, **headers)
41             )
42             try:
43                 next(tiled_data)
44             except StopIteration:
45                 pass
46             except urllib.error.HTTPError:
47                 break
48             for feature in tiled_data:
49                 if fast_comparison:
50                     if lonlat_in_bbox(bbox, feature[0], feature[1]):
51                         yield feature
52                 else:
53                     if shape_contains_feature(sliced, feature):
54                         yield feature

```

Questions