

MVP OSM

**a tool that allows to individuate the areas of high activity
based on level of detail**

Maurizio Napolitano <napo@fbk.eu>

SoNet Group– <http://sonet.fbk.eu>

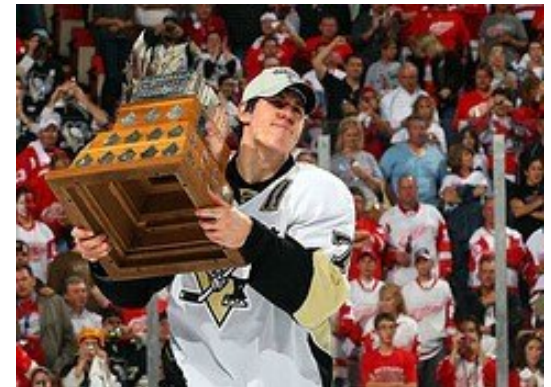


OpenStreetMap like a game

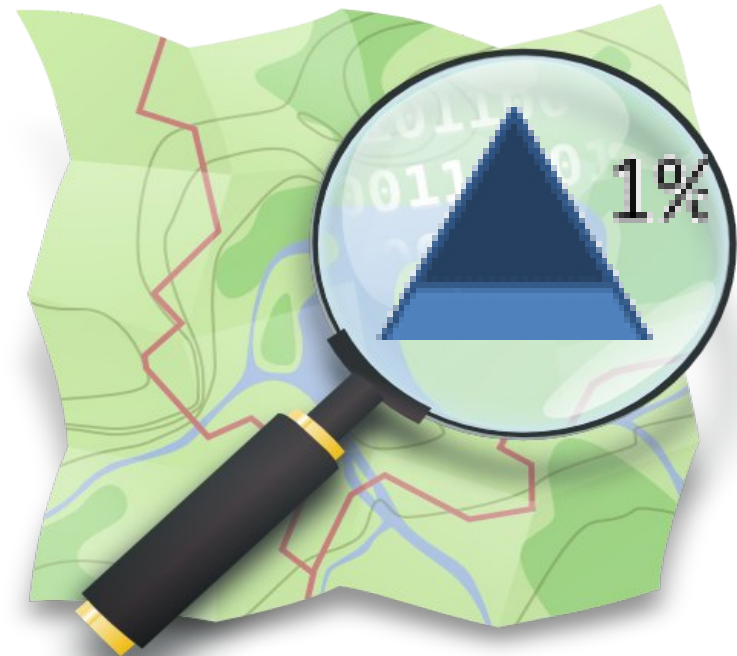
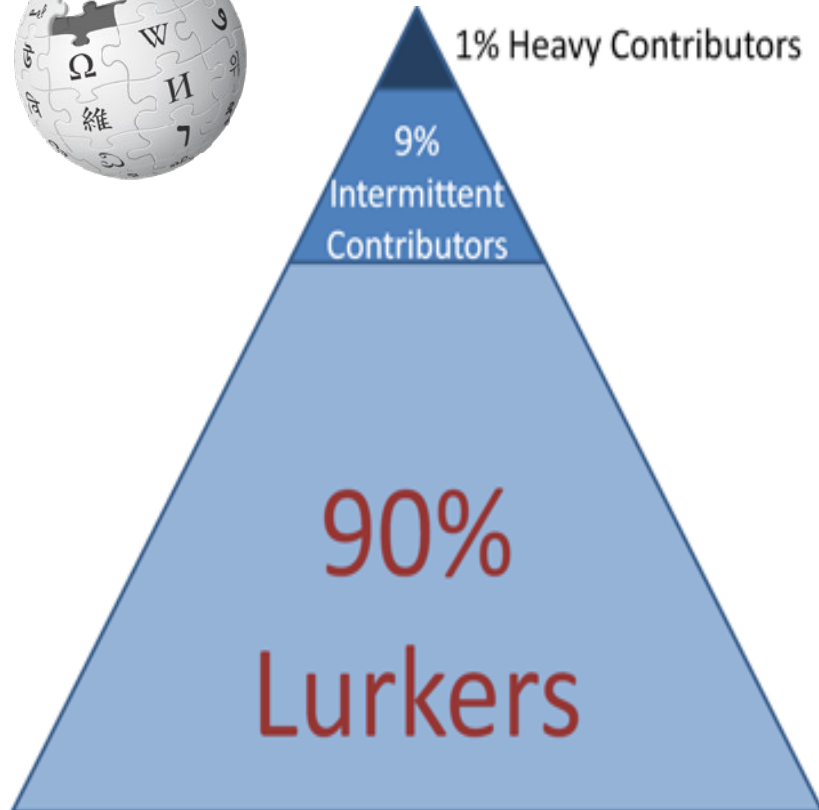


Who are the best players?

Who are the **M**ost **V**aluable **P**layers?



The 90-9-1 Principle



MVP!!!

Jakob Nielsen's Alertbox, October 9, 2006 inShare68
Participation Inequality: Encouraging More Users to Contribute

http://www.useit.com/alertbox/participation_inequality.html 1-9-90

How find?

Common Sense:

a user is a good player if gives many contributions to the map

a test with an italian user



The screenshot shows a web browser window with the address bar displaying <http://www.openstreetmap.org/user/simone>. The page features a navigation bar with buttons for 'Visualizza', 'Modifica', 'Storico', 'Esporta', and 'Tracciati GPS', along with links for 'entra' and 'iscriviti'. The user profile for 'simone' is displayed, including a profile picture of a man with glasses. The profile information states: 'Mappatore dal: 15 Aprile 2005 alle 12:24 (circa 6 anni fa)'. The 'Descrizione' section reads: 'Living near Pavia, north Italy, active on the Italian and international mailinglists. OpenStreetMap Foundation Board Member.' and provides a link to the user's wiki page: <http://wiki.openstreetmap.org/wiki/User:Simone>. The left sidebar contains a search bar and a list of map features.

Nodes created by simone



??
??
??

The complete Italy???

... he is also a “data” importer ...



*[...]
I'm the prime guilty for the
import of
the italian internal borders
between
municipalities, provinces,
regions, and country
[...]*

<http://wiki.openstreetmap.org/wiki/User:Simone#Data>

Solution

Restrict the query only to some details,
those requiring

Local knowledge
Experience
Recognition

A good example:



`step_count=*`

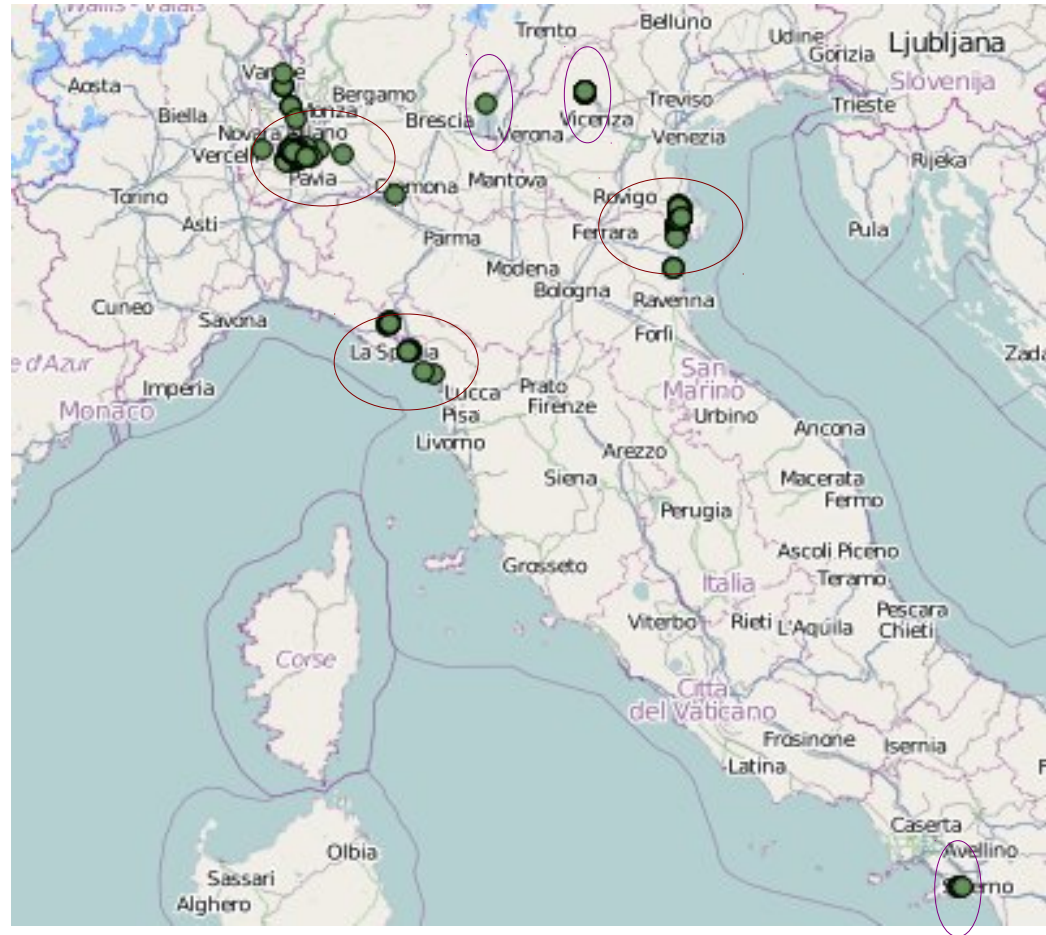
44 keys selected

Keys ...

abandoned
access
access:bicycle
access:bus
access:foot
access:hgv:max_length
access:motorcar
amenity
bicycle
bicycle:backward
dispensing
disused
drinkable
embankment
foot
foot:backward
footway
hiking
horse
incline
oneway

recycling:batteries
recycling:cans
recycling:clothes
recycling:engine_oil
recycling:glass
recycling:paper
recycling:plastic
recycling:plastic_bottles
recycling:plastic_packaging
recycling:scrap_metal
recycling:white_goods
recycling:wood
step.condition
step_count
step.height
step.length
surface
surface.material
traffic_calming
traffic_sign
trail_visibility
visibility

this is the new result



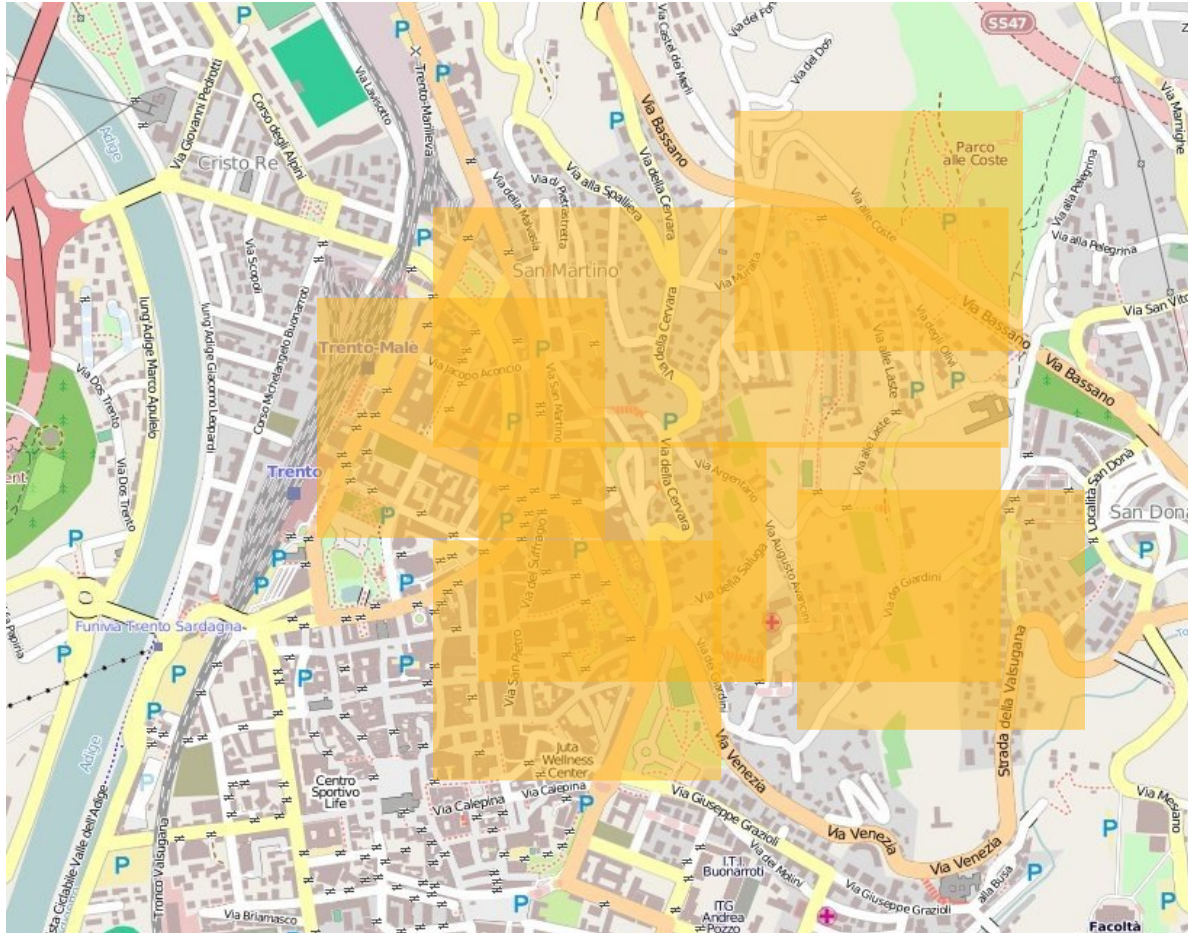
Locations:

Personal life (family, vacancy, work)

OSM life (mapping party, short journeys ...)

The “pet location” concept

have a great care for a location as well as a lovable pet



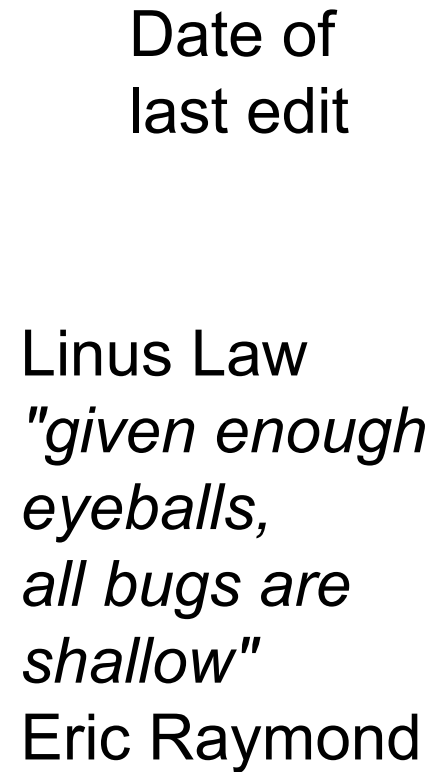
Pet location vs Mapping Party

You will always find a mailbox that isn't mapped

Edoardo Marascalchi
italian osm mapper

frequency of update

...but there is always a little bit of noise



Crowd Quality concept

attempts to quantify the 'collective intelligence of the crowd generating data' in a spatio-temporal context.

User quality

Local knowledge,
Experience and Recognition



Like a MVP player

Feature quality

How many different users contributed to a feature?
How has a feature developed over time?

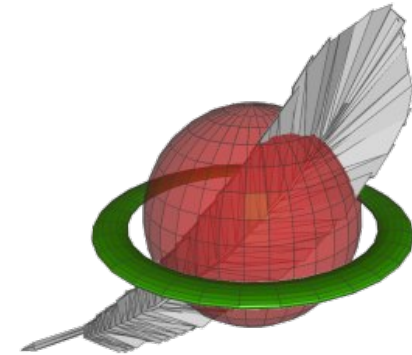
"The impact of crowdsourcing on spatial data quality indicators"
M. van Exel, E. Dias, S. Fruijtjer - 2010

How MVP OSM works

osm xml

```
<?xml version='1.0' encoding='UTF-8'?>
<osm version="0.6" generator="Osmosis.0.32">
  ..<bound box="45.68083,10.37794,47.09520,12.48232" origin="http://
www.openstreetmap.org/api/0.6"/>
  ..<node id="9191867" version="4" timestamp="2010-09-01T22:36:26Z"
uid="330007" user="pikappa79" changeset="5658159" lat="46.104119"
lon="11.0899633"/>
  ..<node id="9191891" version="5" timestamp="2009-03-11T13:30:43Z"
uid="104478" user="giacos73" changeset="786213" lat="46.3693138"
lon="11.290448"/>
  ..<node id="9191920" version="3" timestamp="2009-03-10T12:43:40Z"
uid="104478" user="giacos73" changeset="775964" lat="46.2262194"
lon="11.1514362"/>
```

spatialite



spatialite_osm_raw

python



sql

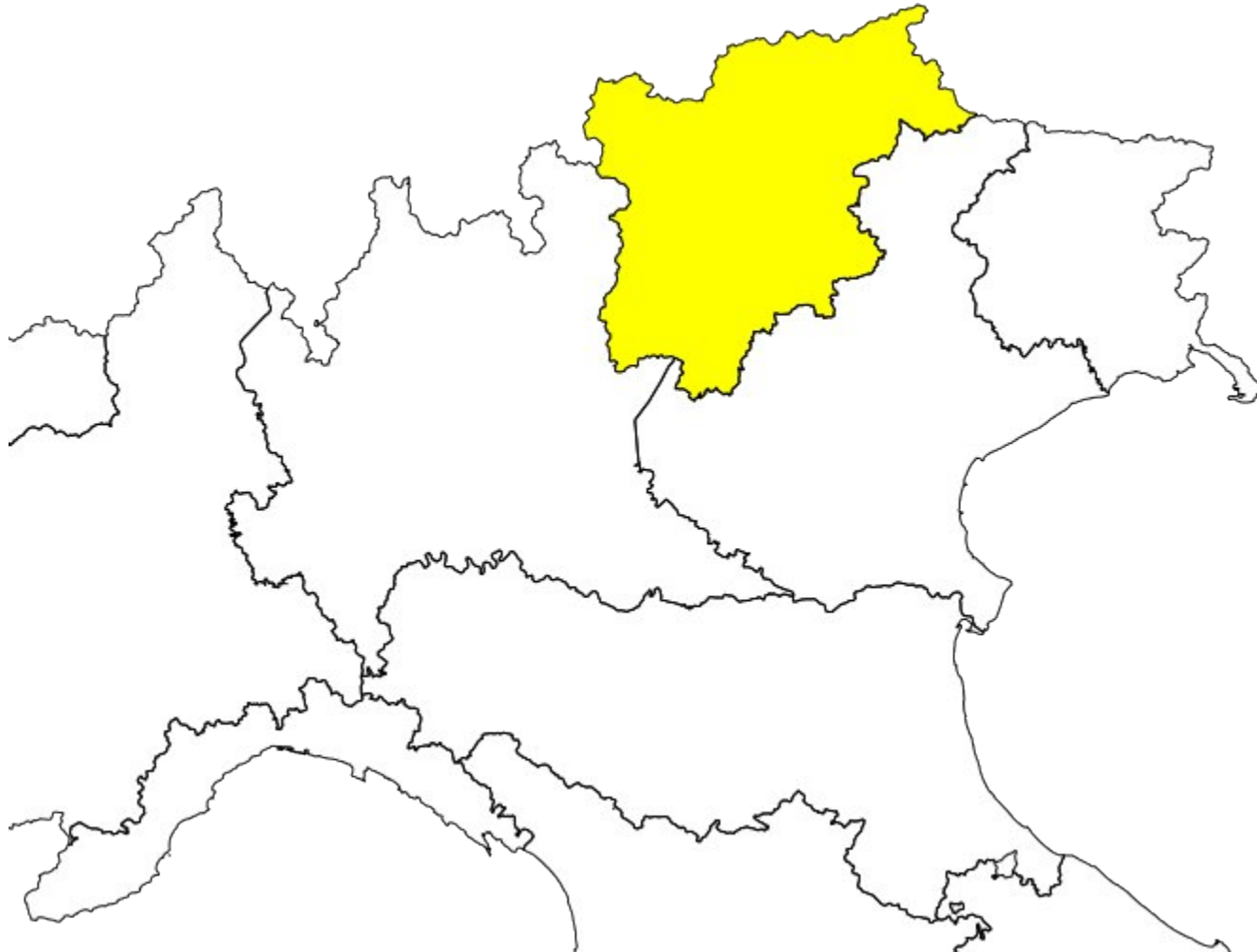
different vectors
for a gis analysis

The imported tables structure

```
spatialite_osm_raw -o map.osm -d database.sqlite
```

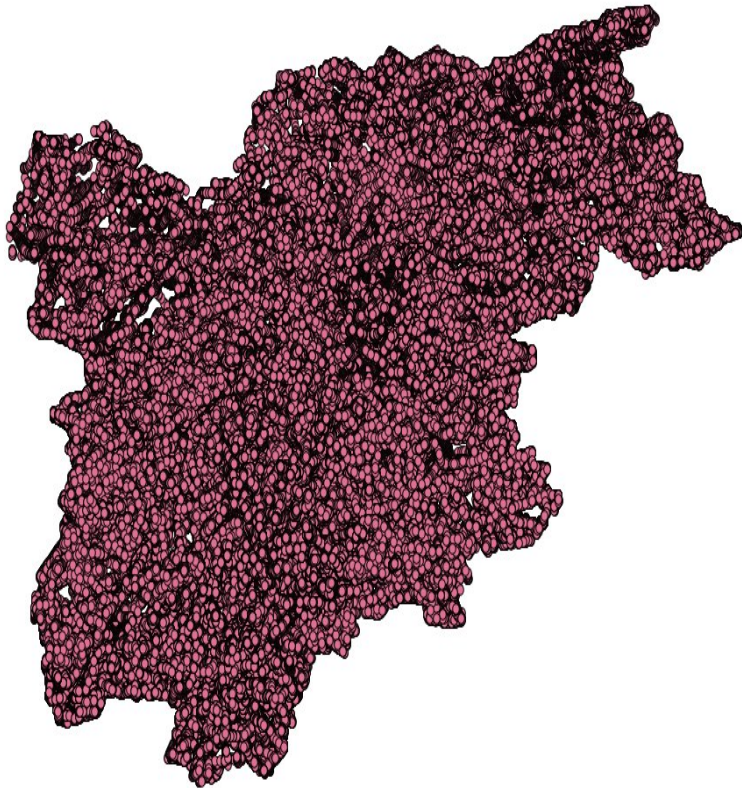
```
osm_relation_way_refs
osm_relations
osm_way_node_refs
osm_node_tags
osm_way_tags
osm_nodes      ───────────► GEOMETRY (points)
osm_ways
osm_relation_node_refs
osm_relation_relation_refs
osm_relation_tags
[...]
```

An example with Trentino Alto Adige

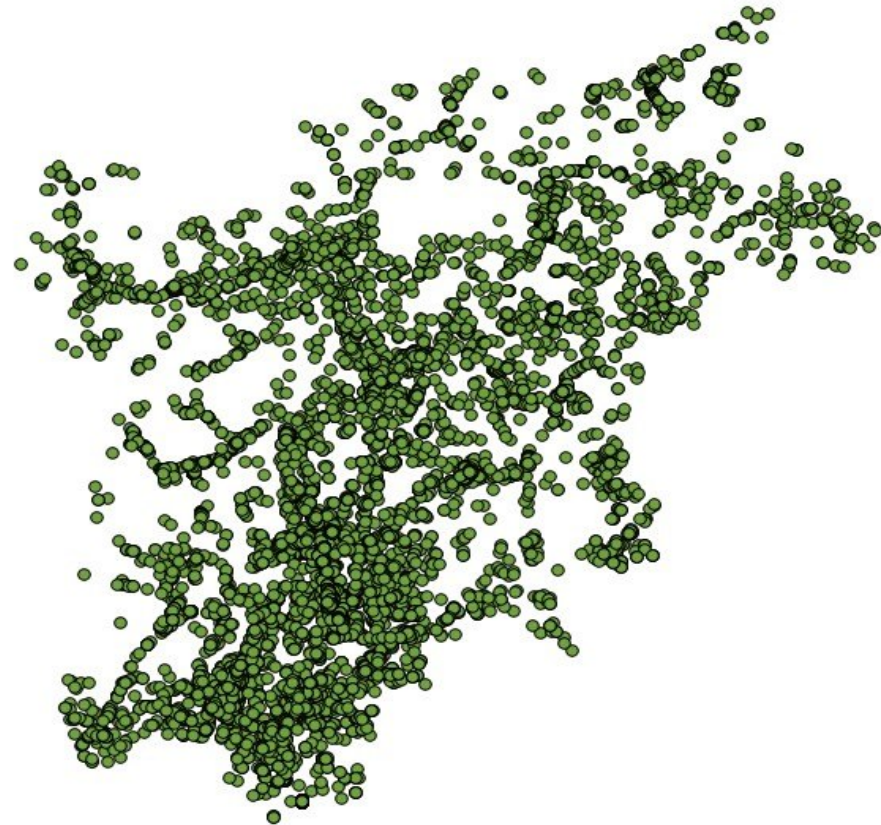


STEP 1/3 – find the points

Query to extract the details



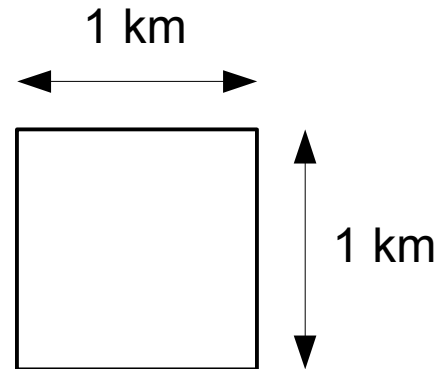
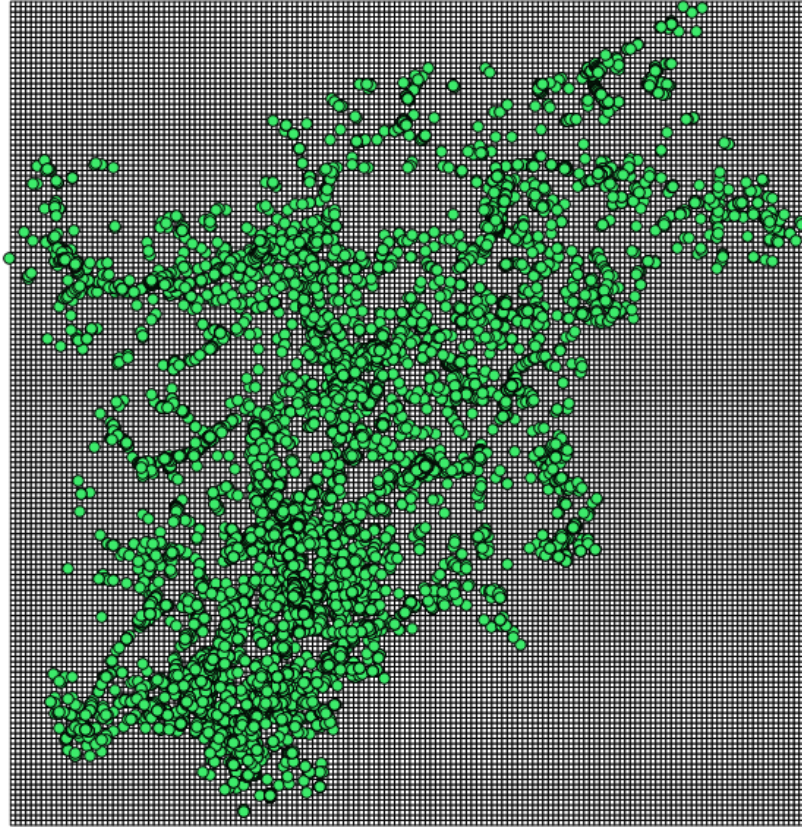
1253 users contributed to this map.
2465872 points



600 users match the
query
13675 points

STEP 2/3 - Density

create a grid ...

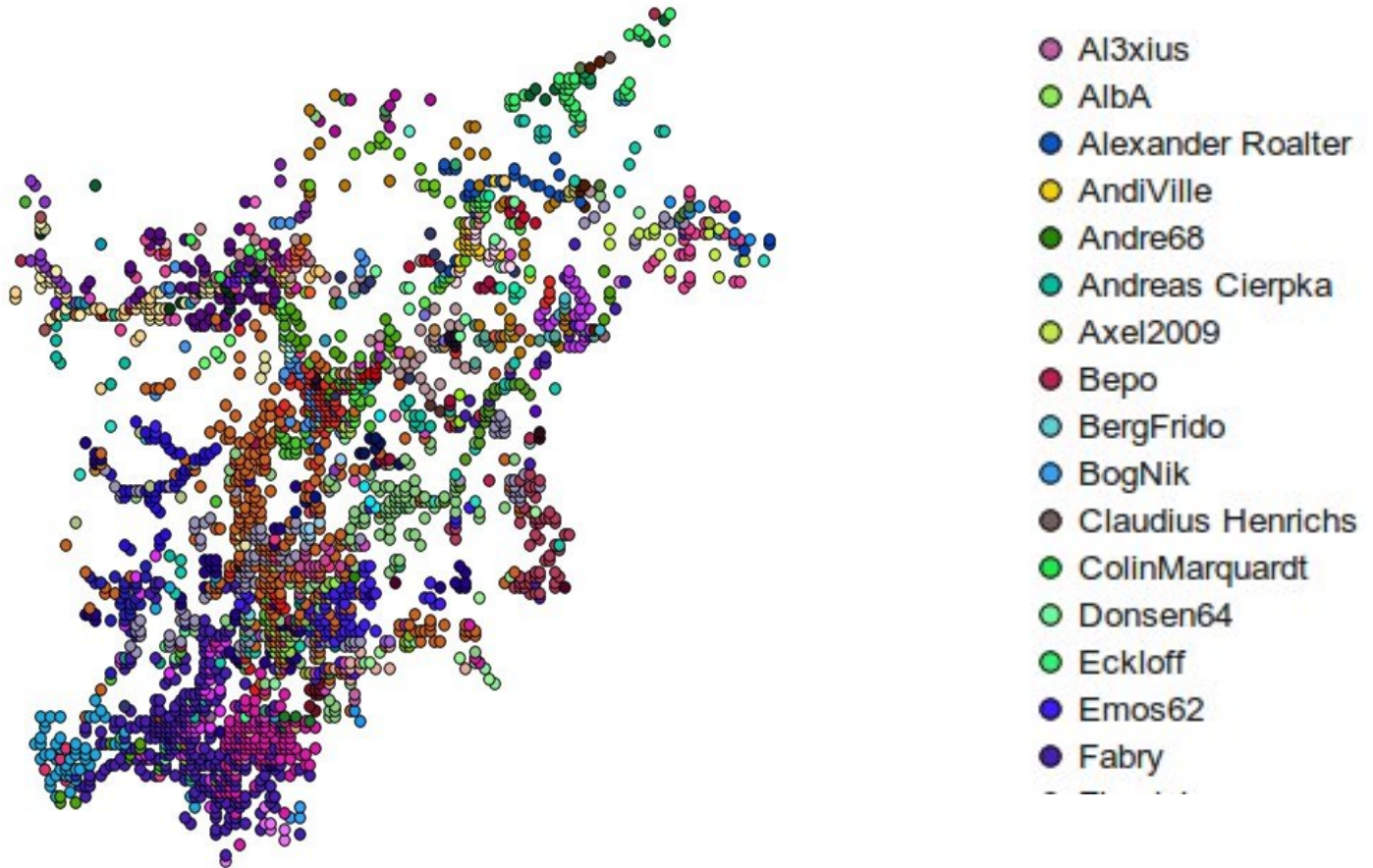


(in this case)

STEP 2/3 - Density

... calculate the data density for each user

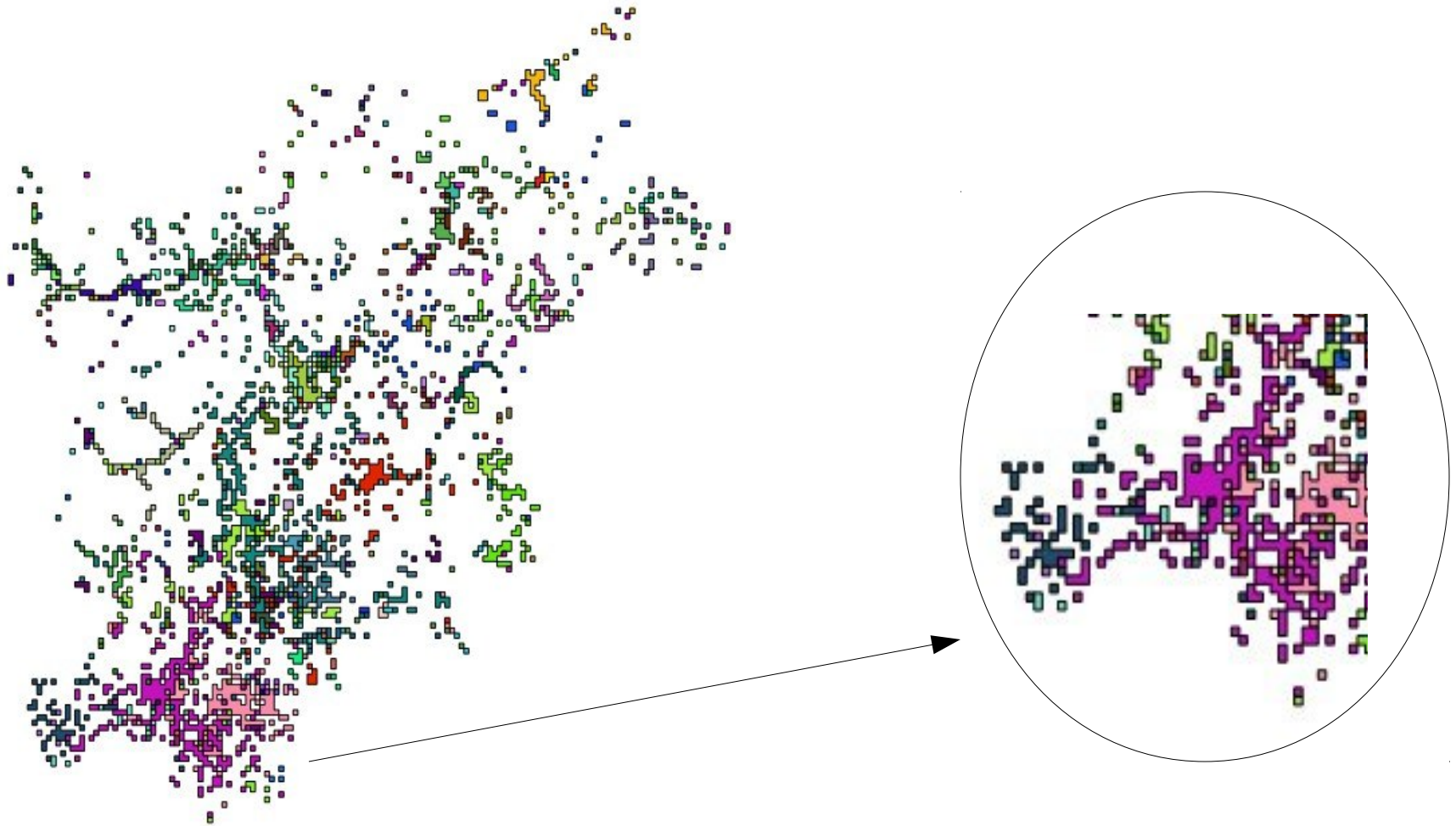
2.444 cells



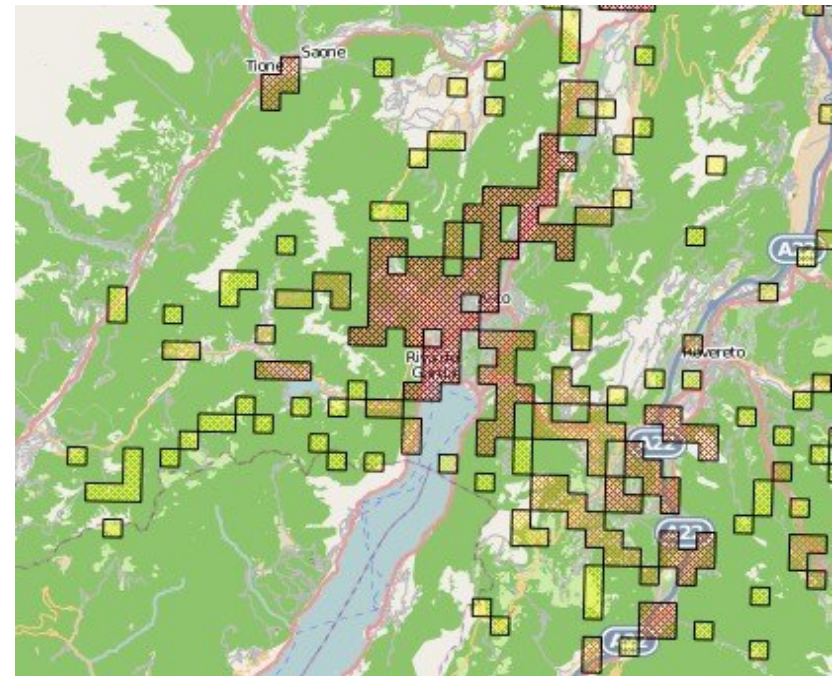
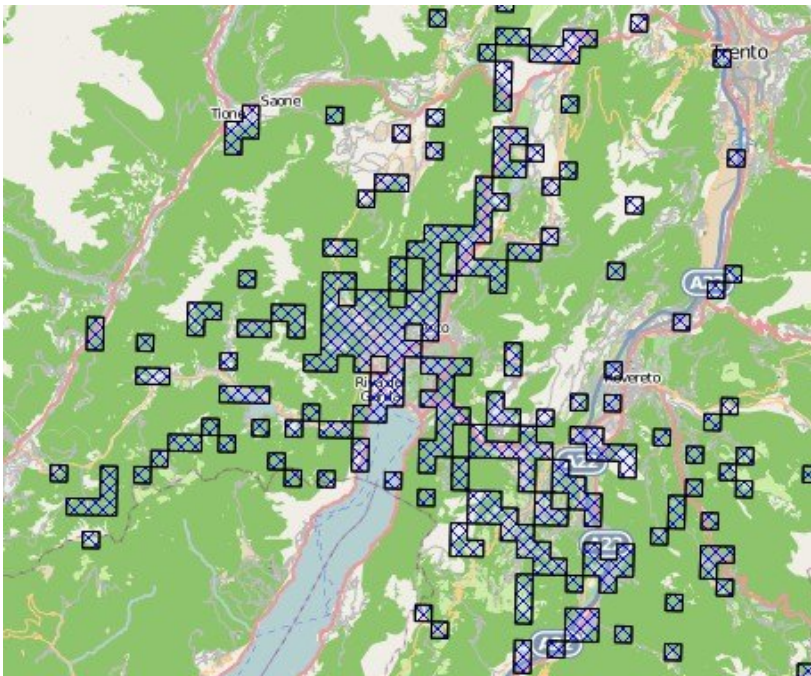
this operation calculate only the users active over the past 3 months
In this case 147 users

STEP 3/3 - cluster

cluster the data for each user by distance

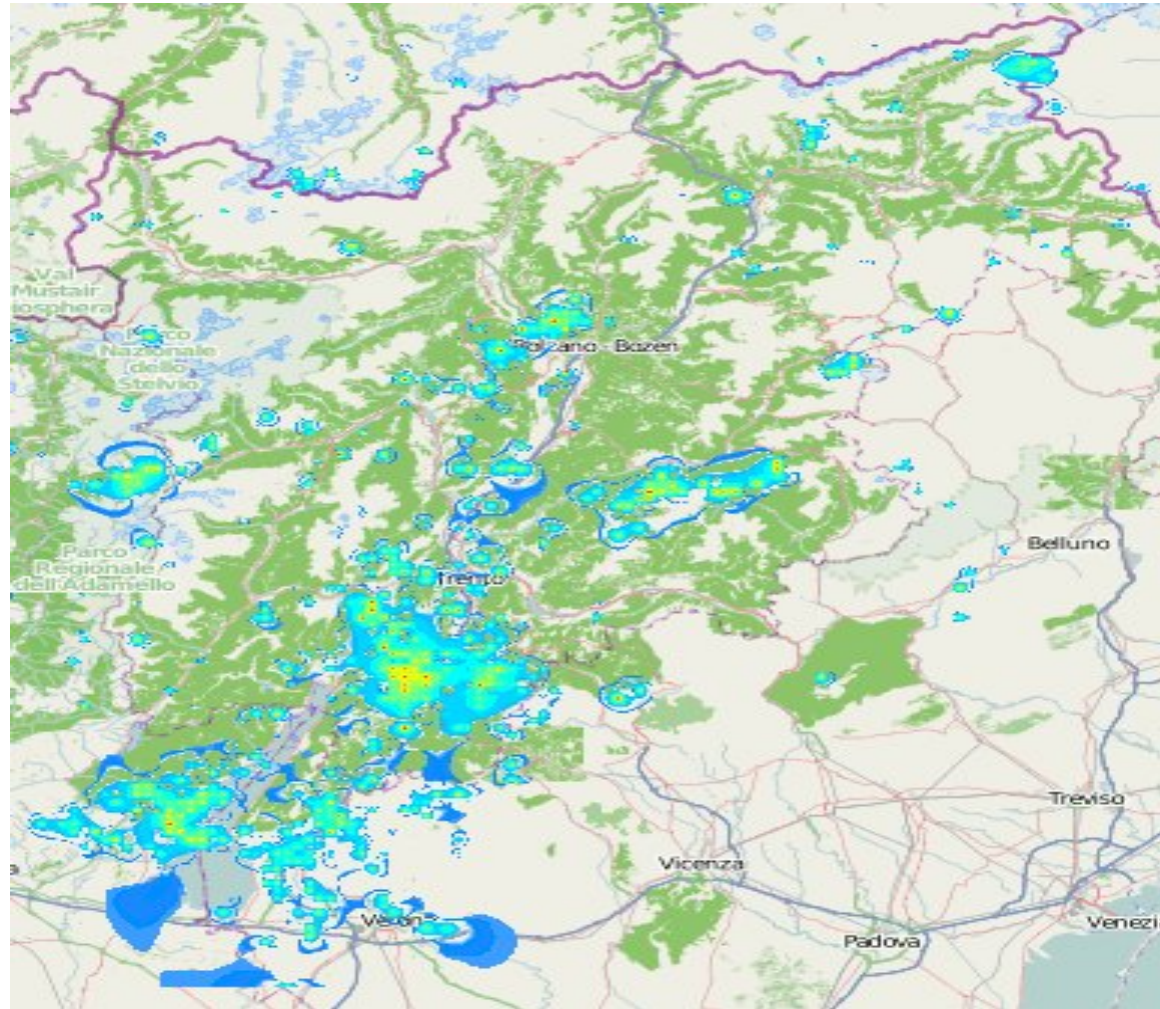


Data Analysis (1/2)



Pet location for a user (1) and where is the most activity (2)

Data analysis 2/2



heat map of the locations

End

Thanks for the attention

Special thanks to Alessandro Furieri
(spatialite head developer)
for the support



Attribution 3.0 Unported