Role of Volunteered Geographic Information in Advancing Science: Effective Utilization

September 23, 2014

POLITECNICO DI MILANO
GEO Laboratory
Como Campus

CC BY NC SA

Free and Open Source Software for citizen science applications

Maria Antonia Brovelli, Marco Minghini, Giorgio Zamboni

Politecnico di Milano, DICA – GEO Laboratory at Como Campus

Citizen science

- ✔ Set of practices in which citizens participate in data collection, analysis and dissemination of a scientific project (Cohn 2008)
- ✔ Active or passive
- ✔ Explicit or implicit
- ✔ Classification (Haklay 2013)
  - ➔ 'classic' citizen science: amateurs engaged in traditional scientific activities
  - ➔ community science: measurements and analysis carried out by amateurs in order to set action plans to deal with environmental problems
  - ➔ citizen cyberscience: use of computers, GPS receivers and mobile phones
    - ✗ volunteered computing: citizens download data, run analyses on their own computers and send back data to the server
    - ✗ volunteered thinking: citizens perform classification works
    - ✗ participatory sensing: applications centered on mobile phones capabilities

Citizen science – Our experience

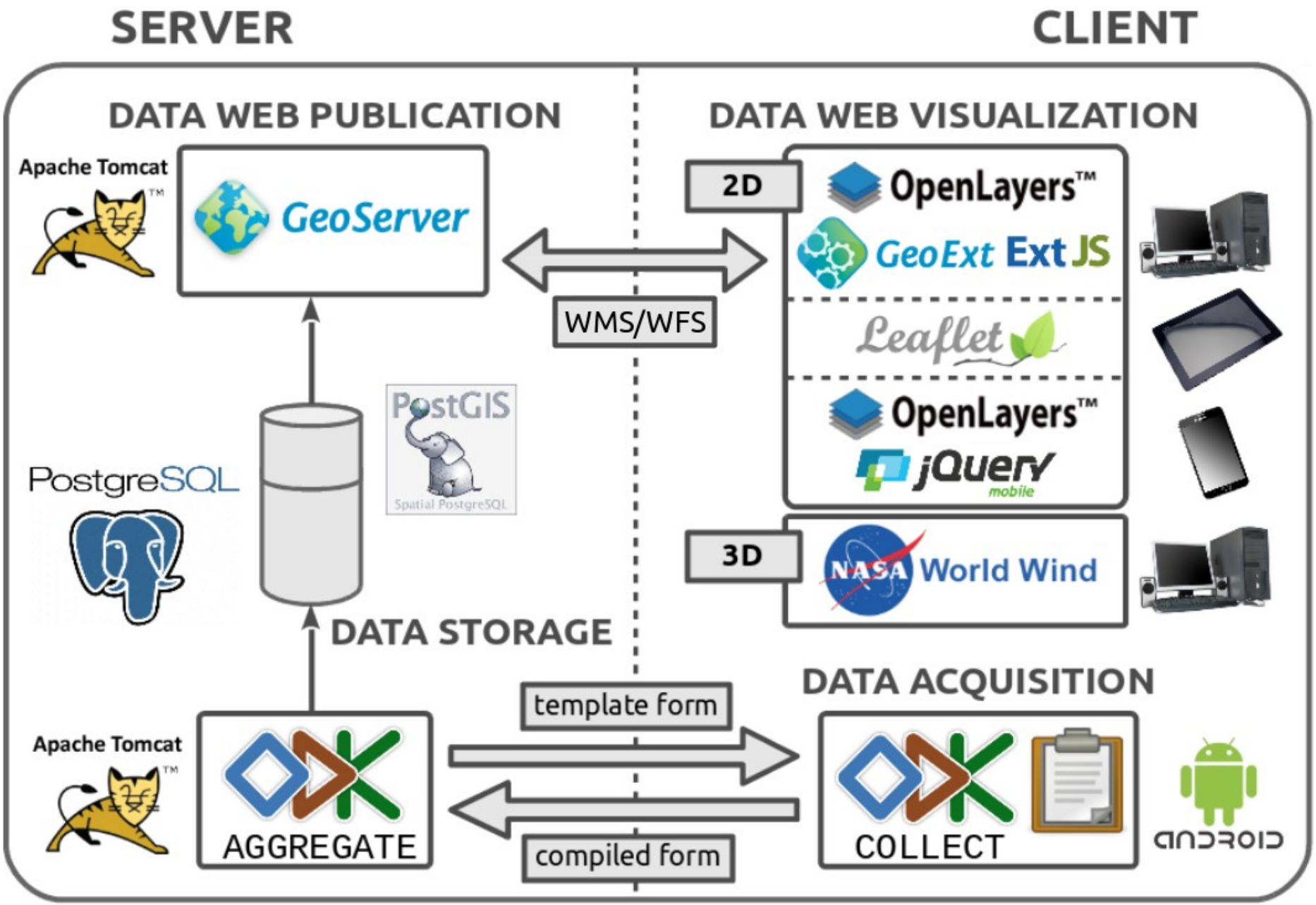
- ✔ Set of practices in which citizens participate in data collection, analysis and dissemination of a scientific project (Cohn 2008)
- ✔ Active or passive
- ✔ Explicit or implicit
- ✔ Classification (Haklay 2013)
  - ➔ 'classic' citizen science: amateurs engaged in traditional scientific activities
  - ➔ community science: measurements and analysis carried out by amateurs in order to set action plans to deal with environmental problems
  - ➔ citizen cyberscience: use of computers, GPS receivers and mobile phones
    - ✗ volunteered computing: citizens download data, run analyses on their own computers and send back data to the server
    - ✗ volunteered thinking: citizens perform classification works
    - ✗ participatory sensing: applications centered on mobile phones capabilities

Our experience

- To build an architecture for participatory sensing applications enabling:
  - field-data collection from mobile devices
  - data access and interaction based on user privileges
  - data Web publication and sharing
  - data access through multi-dimensional (2D and 3D) Web interfaces
  - data participative enrichment

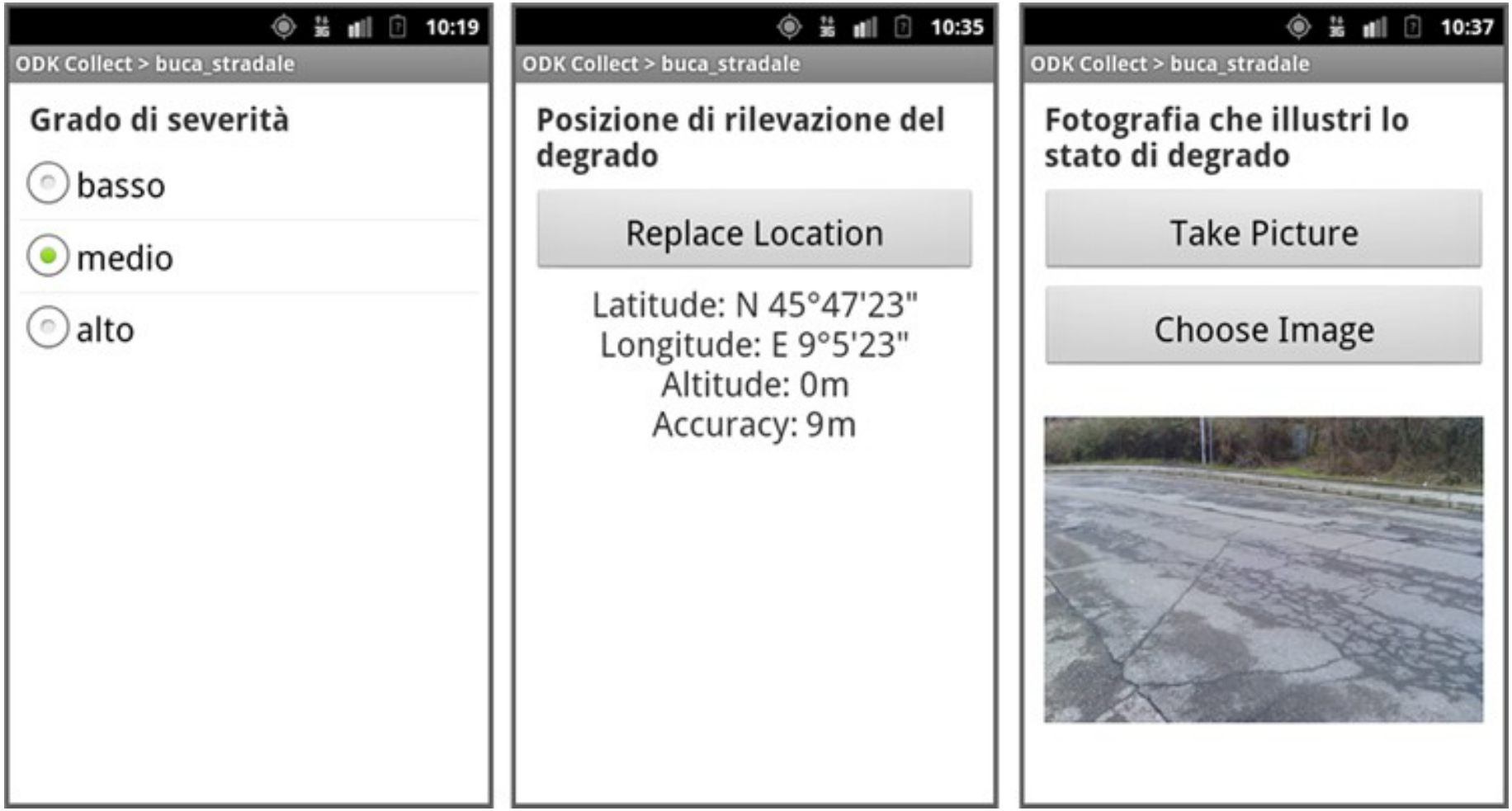
- To use FOSS (Free and Open Source Software)
  - Involvement in OSGeo (Open Source For Geospatial)
  - Involvement in GeoForAll
    ICA-OSGeo-ISPRS initiative http://www.geoforall.org/
    89 Labs as of 20 September 2014

Participatory sensing FOSS architecture



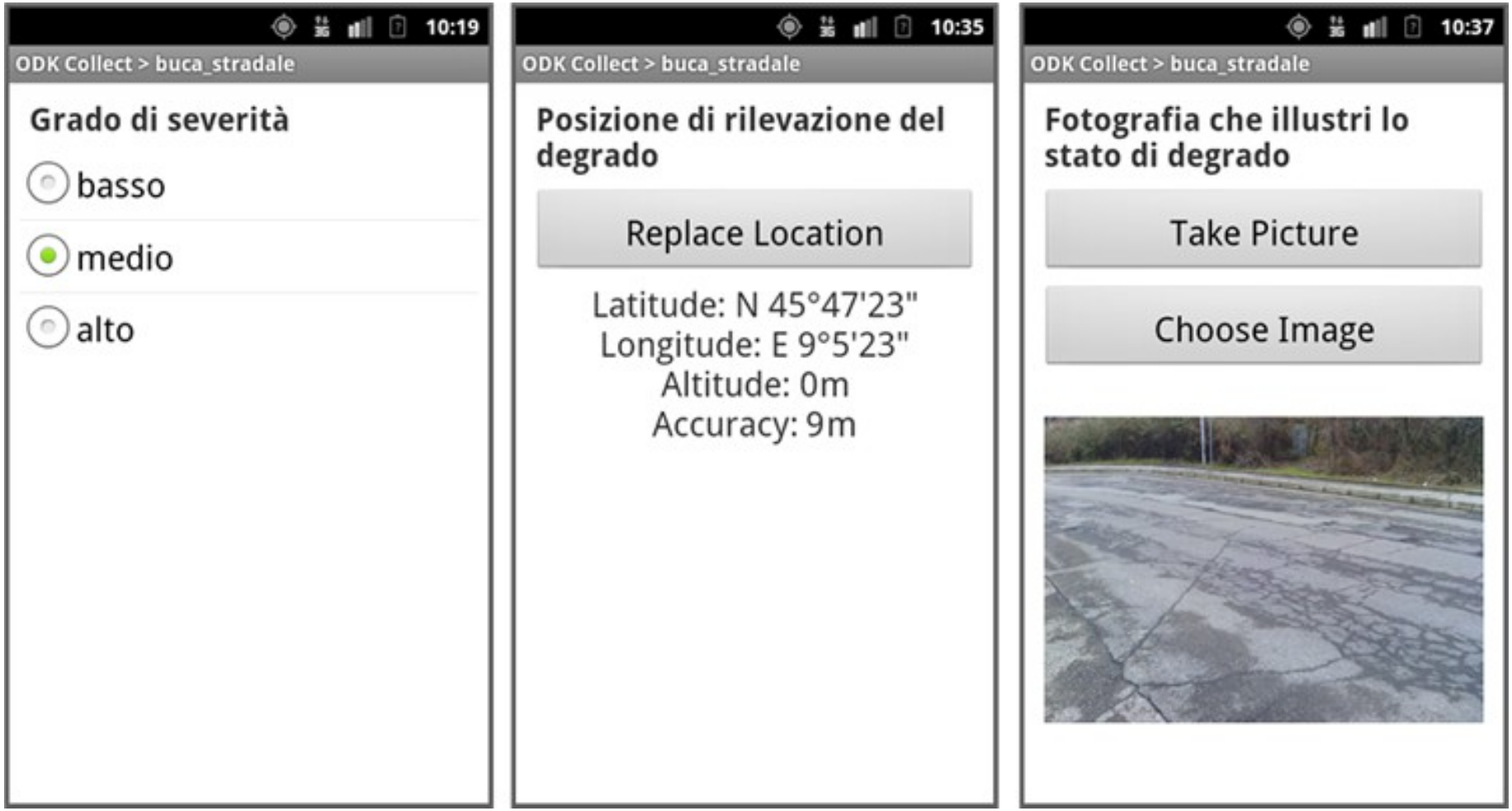
Report of potholes

- Report of road pavement damages, winter 2012-13 (Brovelli et al. 2014a)
  - classification of the type of damage and its severity degree
  - open to everyone, advertised through local media (TV and newspapers)

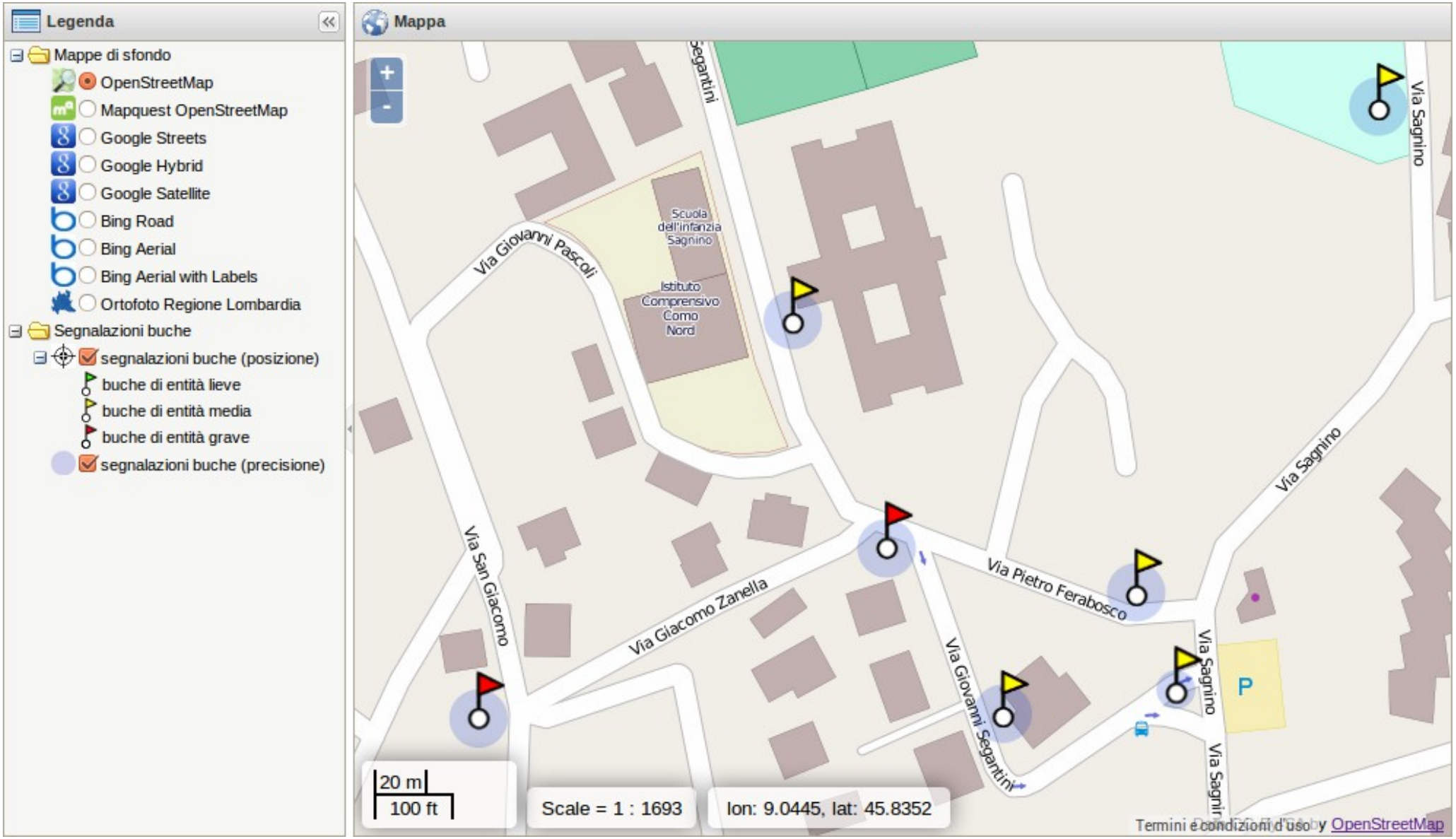


Report of potholes

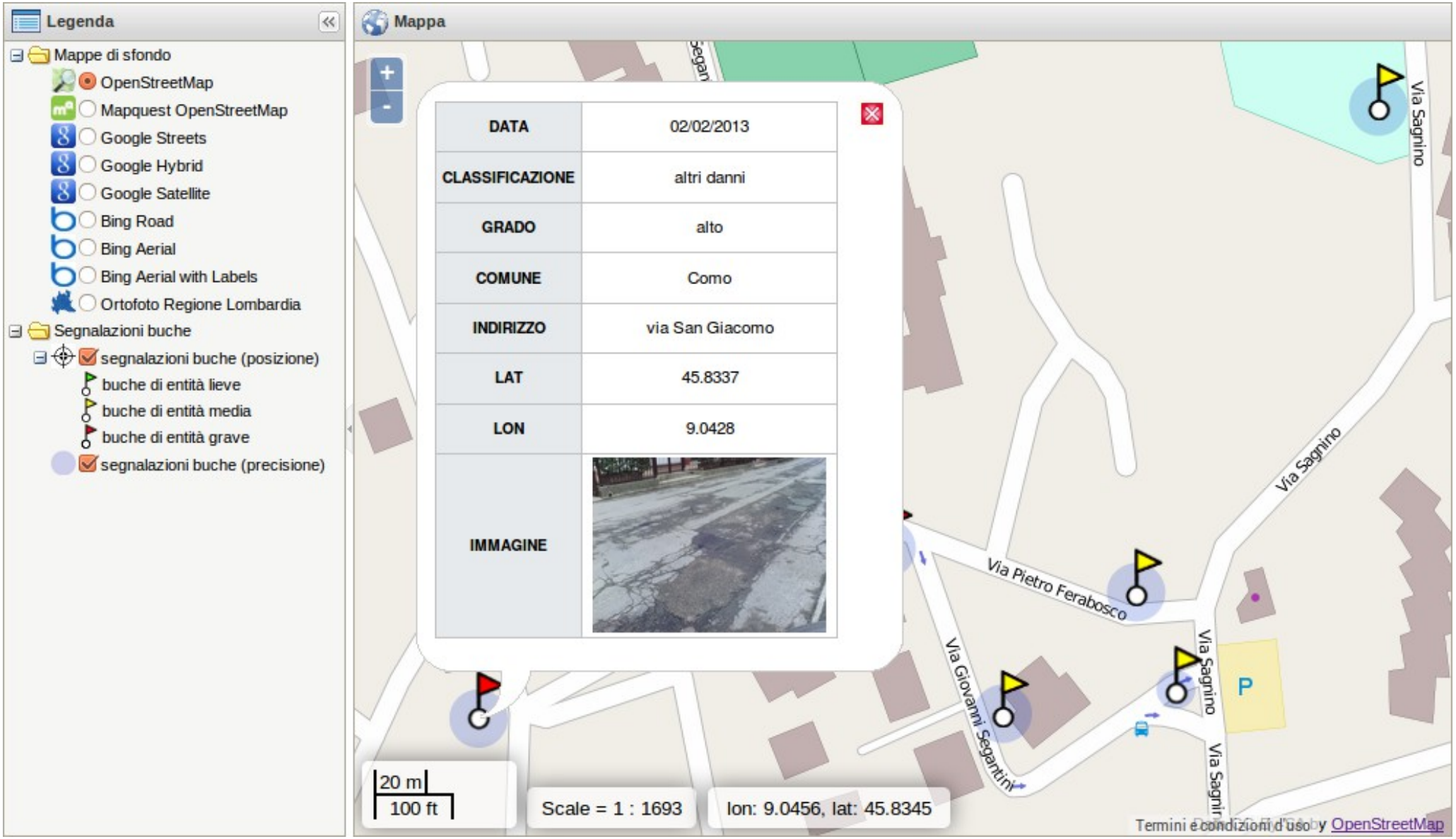
- Report of road pavement damages, winter 2012-13 (Brovelli et al. 2014a)
  - classification of the type of damage and its severity degree
  - open to everyone, advertised through local media (TV and newspapers)



Report of potholes

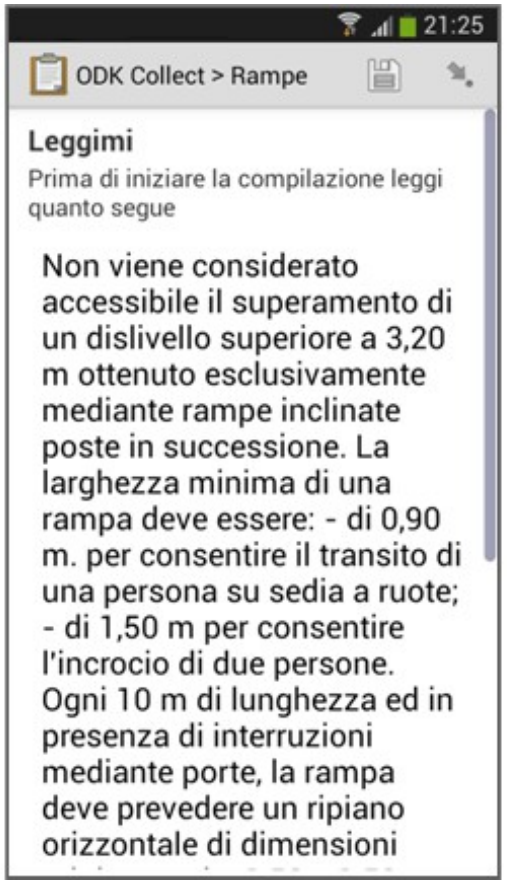


Report of potholes



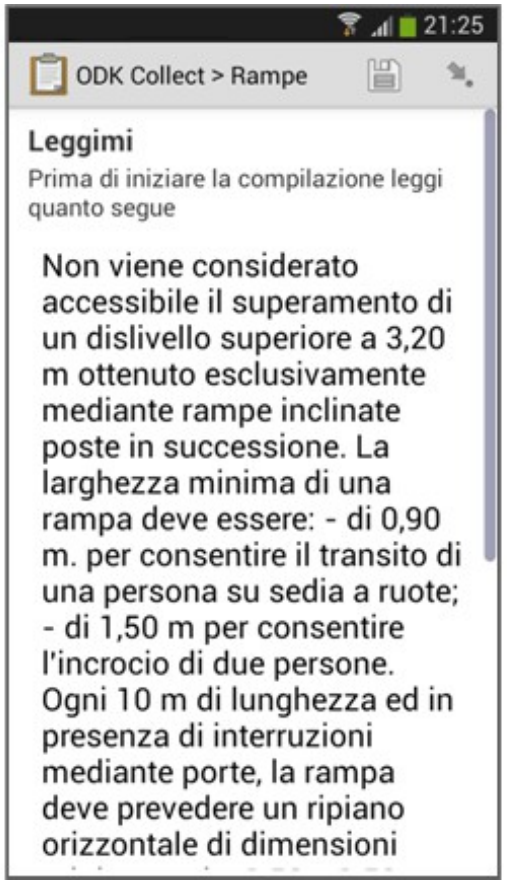
Report of architectural barriers

- Report of architectural barriers, fall 2013
  - physical obstacles for the mobility of people (stairs, ramps, pathways)
  - verification of the compliance with the current Italian legislation
  - performed by students within a collaboration action with a local high school

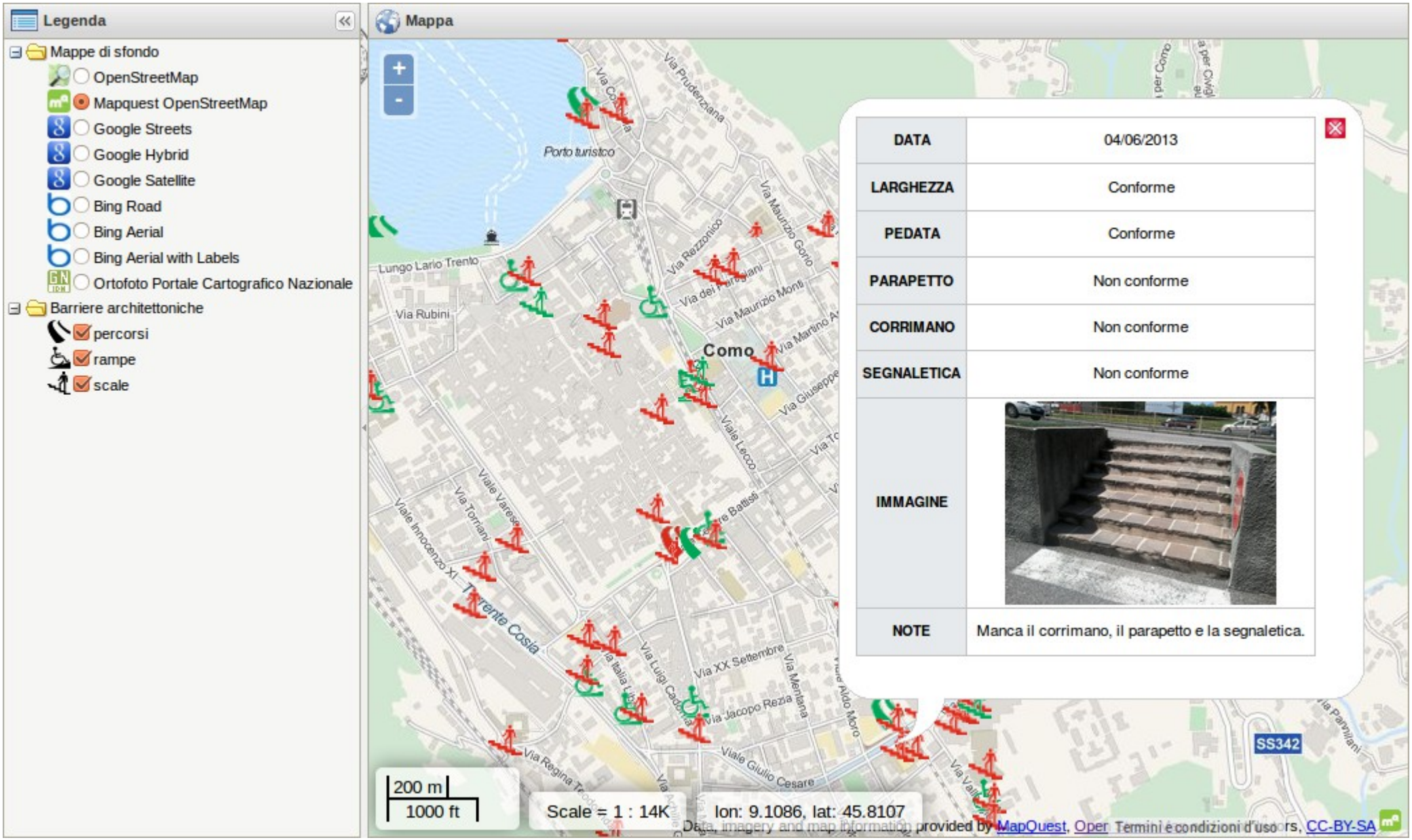


Report of architectural barriers

- Report of architectural barriers, fall 2013
  - obstacles for physically impaired people (stairs, ramps, pathways)
  - verification of the compliance with the current Italian legislation
  - performed by students within a collaboration action with a local high school



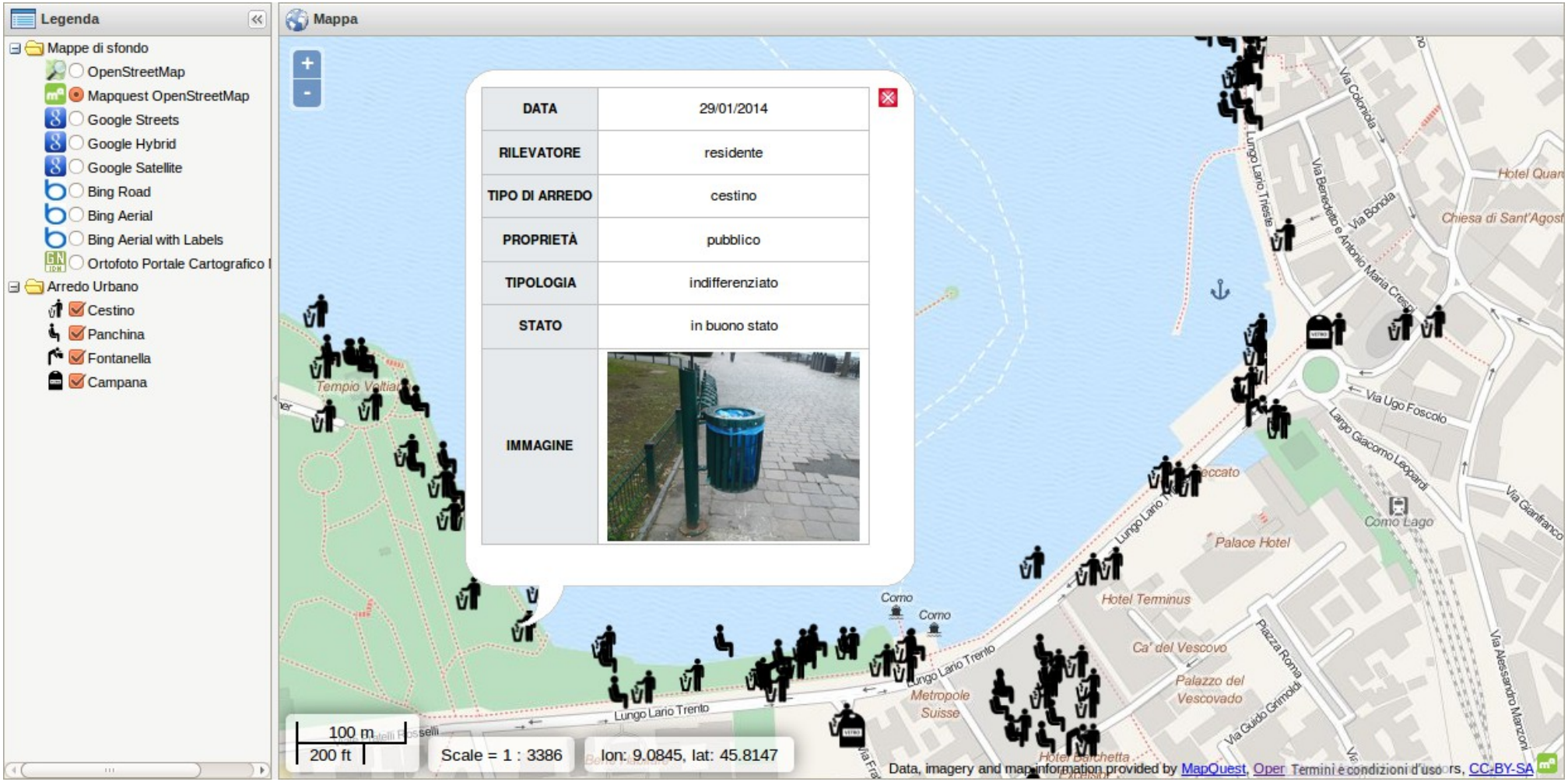
Report of architectural barriers



DATA	04/06/2013
LARGHEZZA	Conforme
PEDATA	Conforme
PARAPETTO	Non conforme
CORRIMANO	Non conforme
SEGNALETICA	Non conforme
IMMAGINE	
NOTE	Manca il corrimano, il parapetto e la segnaletica.

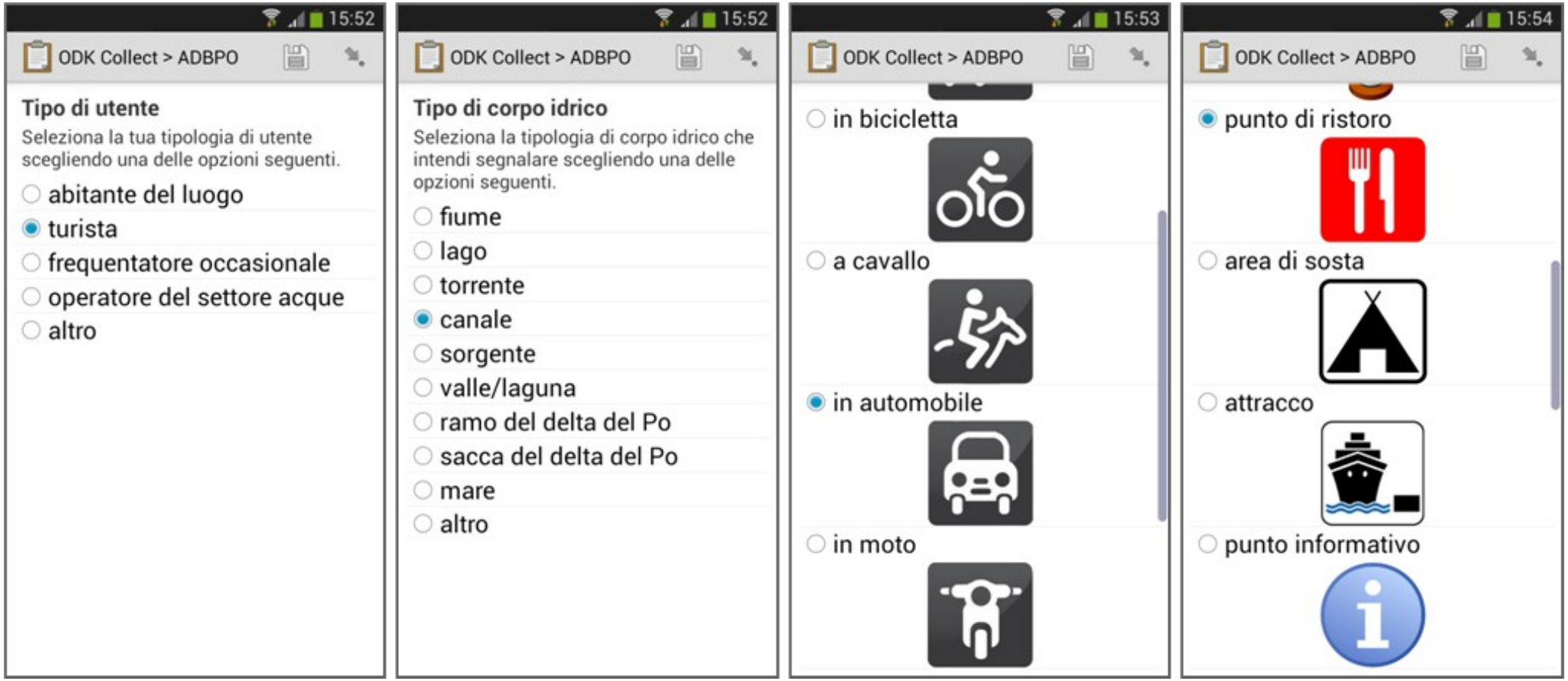
Mapping street furniture

- ✔ Report of street furniture elements, spring 2014
  - ➔ benches, litter bins, waste bells, fountains
  - ➔ performed by students within a collaboration action with a local high school



AdbPo report of water-related elements

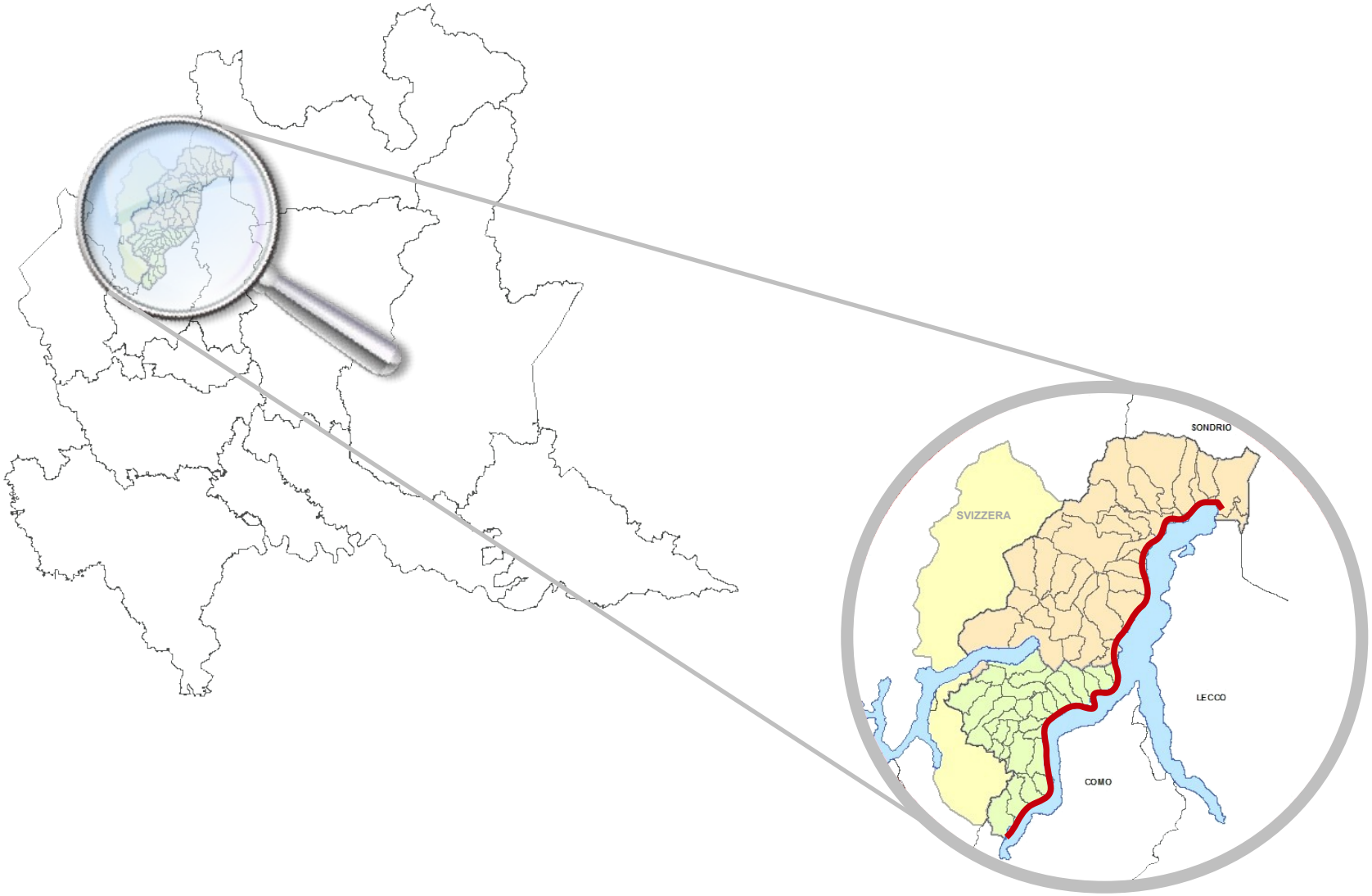
- ✔ Report of water-related elements, ongoing (Brovelli et al. 2014b)
  - ➔ VGI campaign managed by AdbPo (Autorità di Bacino del fiume Po – *Po river basin Authority*) to enlarge consciousness of its territory through collection and valorisation of user local knowledge and experience



AdbPo report of water-related elements

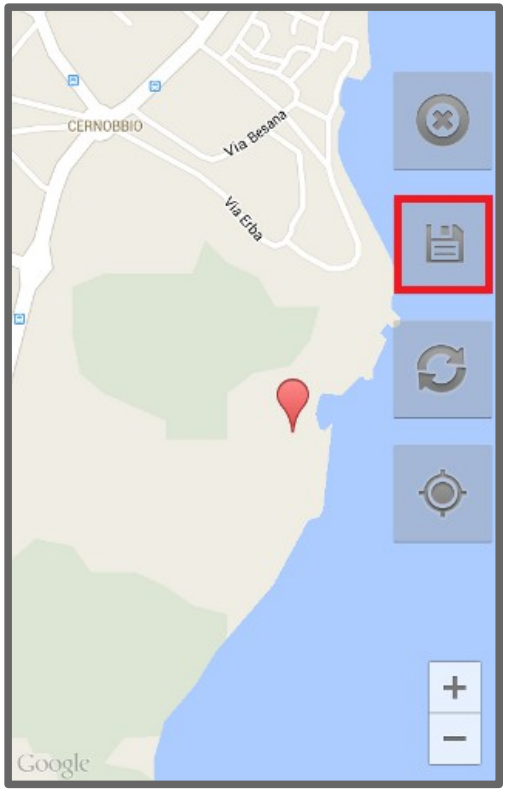
Mapping elements along cultural paths

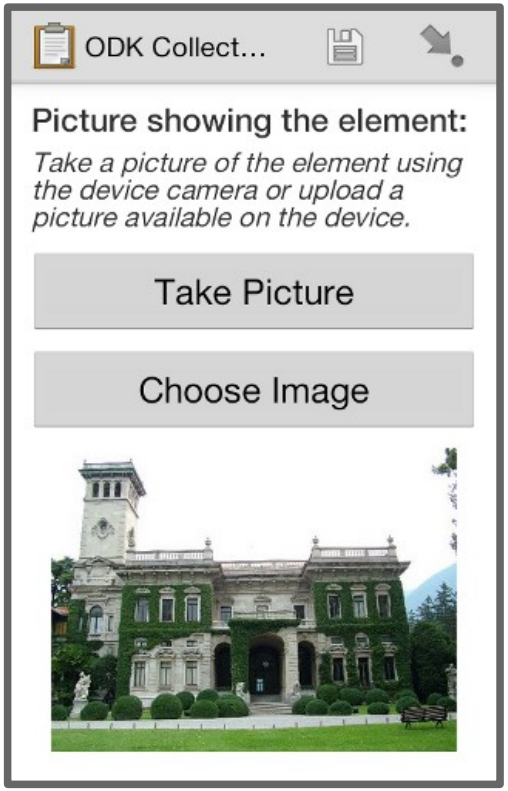
- ✔ Map parties within INTERREG project “The paths of Via Regina”, ongoing



Mapping elements along cultural paths

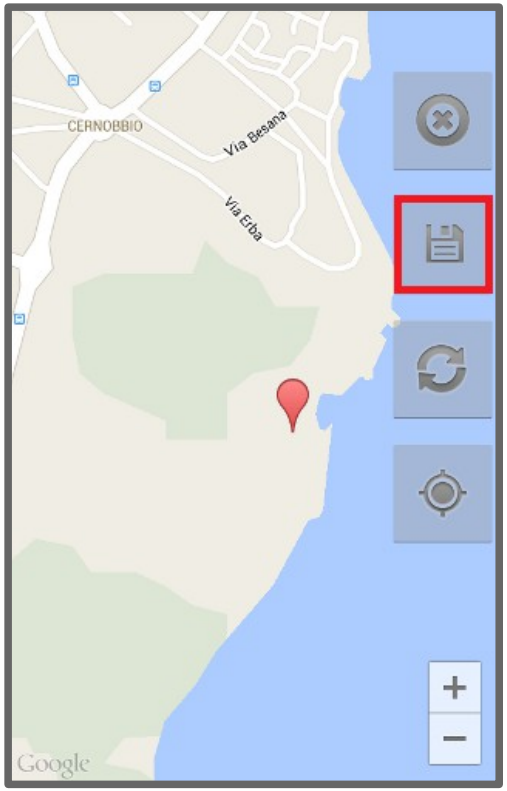
- ✔ Map parties within INTERREG project “The paths of Via Regina”, ongoing
  - user report of points of interest related to history, culture, tourism and other services along the paths connected to the ancient Via Regina
  - one-day events (jointly organized with a cultural association of hikers) which target local communities and are advertised ad hoc

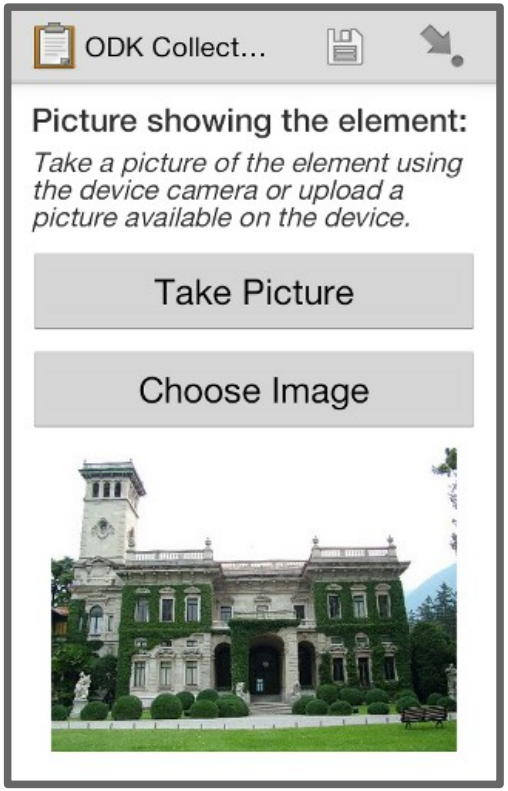




Mapping elements along cultural paths

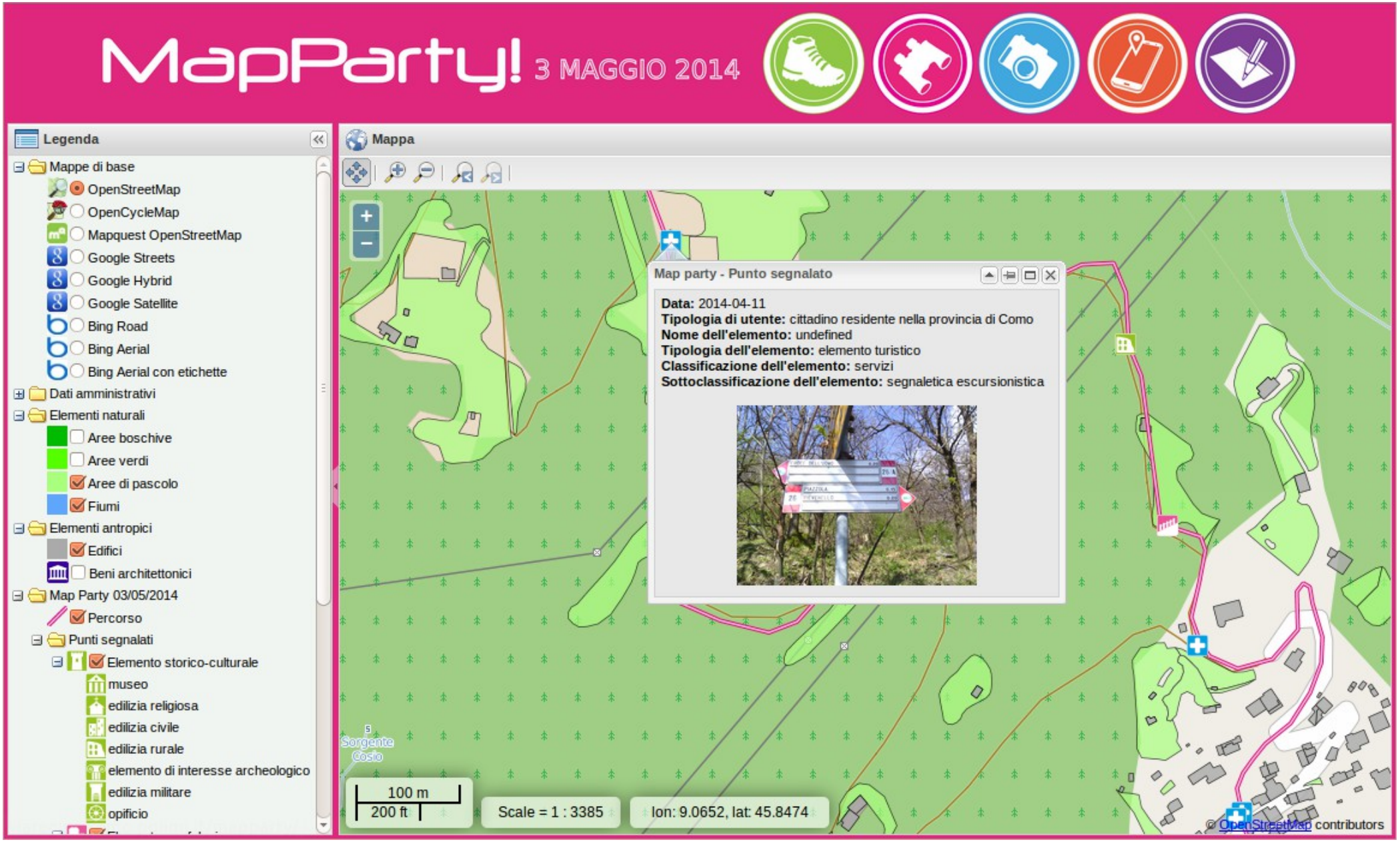
- ✔ Map parties within INTERREG project “The paths of Via Regina”, ongoing
  - user report of points of interest related to history, culture, tourism and other services along the paths connected to the ancient Via Regina
  - one-day events (jointly organized with a cultural association of hikers) which target local communities and are advertised ad hoc





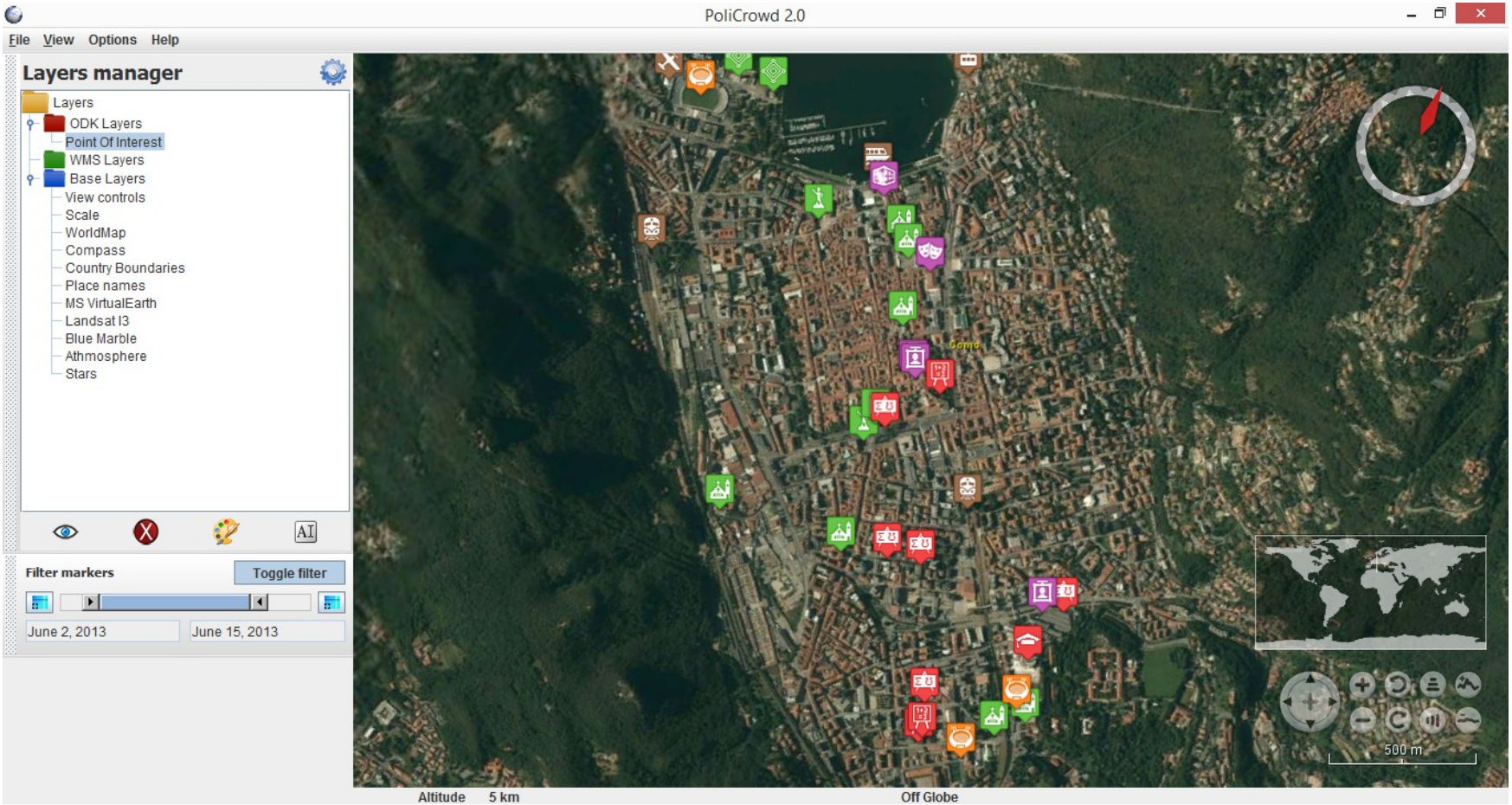
Mapping elements along cultural paths

Mapping elements along cultural paths



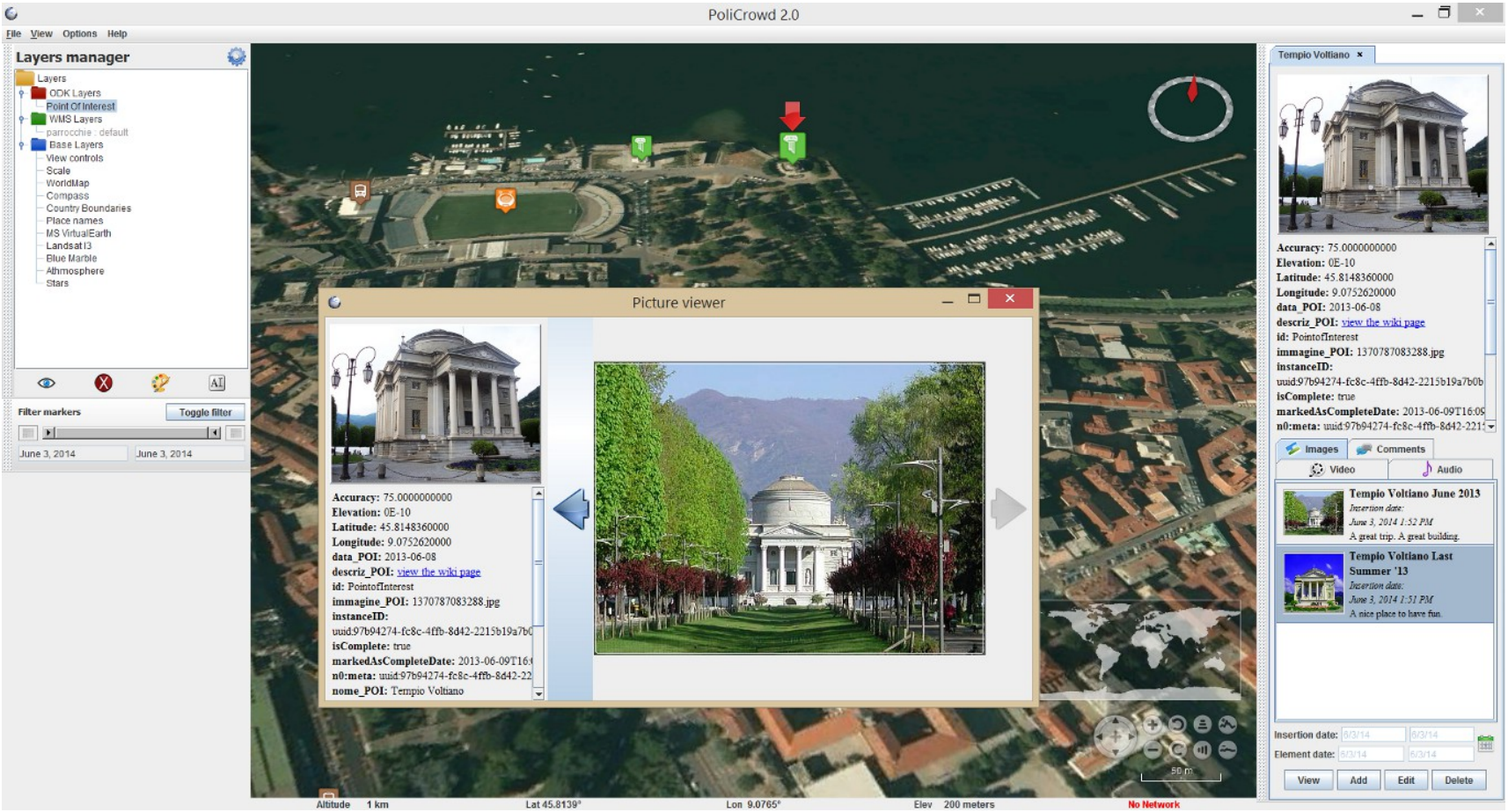
PoliCrowd – A social World Wind platform

- Web-based 3D participatory platform, ongoing (Brovelli et al. 2013)
  - Born for tourism, culture, sports & transportation Points Of Interest (POIs)
  - POIs 3D visualization on NASA World Wind virtual globe
  - participative functionalities: POIs collaborative enrichment & project creation



PoliCrowd2.0 – A social World Wind platform

- Web-based 3D participatory platform, ongoing (Brovelli et al. 2013)
  - non mono-thematic application – connection to any WMS and ODK server
  - customizable data styling & multimedia support
  - support for time dimension (4D visualization)



PoliCrowd and Policrowd2.0

- Web-based 3D participatory platform, ongoing (Brovelli et al. 2013)
  - partially developed by PoliMi students (Giorgio Zamboni tutor)
  - winner of the 1st and the 2nd NASA World Wind Europa Challenge

PoliCrowd 2.0
A Social World Wind Platform

http://geomobile.como.polimi.it/policrowd2.0

Results & considerations

- ✔ VGI data quantity
  - ➔ depends on the number of expected users
    - ✗ campaigns open to the general public usually do not foster a widespread participation → many data collected by relatively few users
    - ✗ campaigns targeted to specific user groups turn out to be more effective → many data collected by almost all the users involved
  - ➔ depends on the kind of expected users
    - ✗ the younger and more technology-oriented they are, the more they will participate
  - ➔ depends on the foreseen duration of the campaign
    - ✗ (without incentives or prizes) people are little stimulated to participate in the long time → students, but also adults!
    - ✗ short VGI campaigns (like map parties) are the most effective solution
- ✔ VGI data quality (semantic & logical accuracy, consistency, completeness)
  - ➔ generally very high (much more than expected!)

Comments, ideas & open questions

- ✔ Technical features of VGI platforms
  - FOSS (Free and Open Source Software) is a valid and mature tool to build up complex architectures tailored to the needs
  - open standards (e.g. from W3C and OGC) should be taken as a requirement when dealing with VGI data
- ✔ VGI validation
  - very hard to define general protocols
  - human validation still needed
- ✔ VGI metadata and geocatalogs
  - no standards yet exist
  - which information do we need with VGI?
    - ✗ positional accuracy → included within VGI from mobile device
    - ✗ up-to-dateness → included within VGI from mobile device
    - ✗ type/name of user? what else?

Chairpeople questions

- How do you see the uprising of 'Big Data' era influencing the way VGI is created and used and what are opportunities/challenges in that context?
- How does one generate future requirements for pervasive VGI data creation and on-the-fly processing/analysis?
- What new methods and standards are needed for incorporating VGI in scientific analysis?
- How do we address the data life cycle for VGI?

More than answers ... many doubts

- The increasing number of crowdsourcing platforms will bring us to Big Data. Is it enough of dealing with the current geo-standards? E.g., could be WFS suitable for describing data? is the current ISO 19115 metadata standard suitable?

- Crowdsourced data is not recognized the same quality as the authoritative ones. Are therefore the quality indexes usually used for GI enough? Is it necessary to have also information about the volunteer who contributes the data? How can we contextualize from the social-economical perspective the single user or the community of citizens participating to the collaborative collection? Is it important or not who are the volunteers?

- Volunteers are not generally interested in maintaining their data (unless they are environmental or social organizations). Who will do that? Public agencies? Private companies? Who are the owners of the data and which are the rights they can claim?

Contacts

Thanks for your attention!

Maria Antonia Brovelli

Politecnico di Milano
Laboratorio di Geomatica – Polo Territoriale di Como
Via Valleggio 11, 22100 Como (Italy)

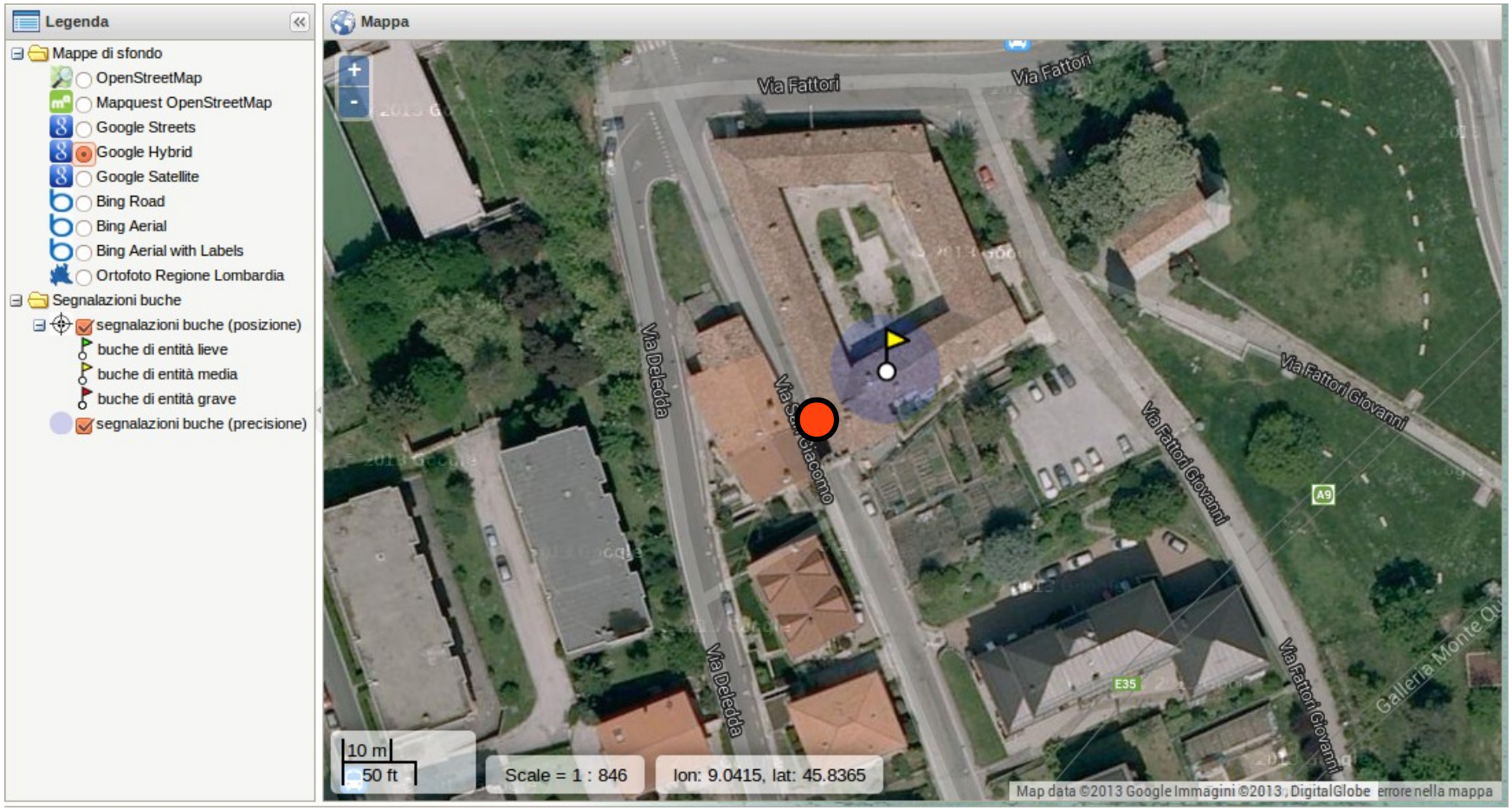
maria.brovelli@polimi.it

References

- Cohn, J. P. (2008). Citizen science: Can volunteers do real research?. *BioScience* 58(3), 192–197.
- Haklay, M. (2013). Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. *In* D. Sui, S. Elwood and M. Goodchild (eds) Crowdsourcing Geographic Knowledge. Dordrecht: Springer. 105–122.
- GNU Project (1996). What is free software? the free software definition [available online] http://www.gnu.org/philosophy/free-sw.en.html (accessed: 2014-09-18).
- Brovelli, M. A., Minghini, M. and Zamboni, G. (2013). Participatory GIS: Experimentations for a 3D social virtual globe. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-2/W2, 13–18.
- Brovelli, M. A., Minghini, M. and Zamboni, G. (2014a). Public Participation GIS: a FOSS architecture enabling field-data collection. *International Journal of Digital Earth* (ahead of print) 1–19. doi: 10.1080/17538947.2014.887150.
- Brovelli, M. A., Dotti, L., Minghini, M., Pancaldi, M. and Zamboni, G. (2014b). Volunteered Geographic Information for water management: a prototype architecture. Proceedings of th 11th International Conference on Hydroinformatics, New York, August 17-21, 2014.
- Minghini, M. (2014). Multi-dimensional GeoWeb platforms for citizen science and civic engagement applications. PhD thesis, Politecnico di Milano.

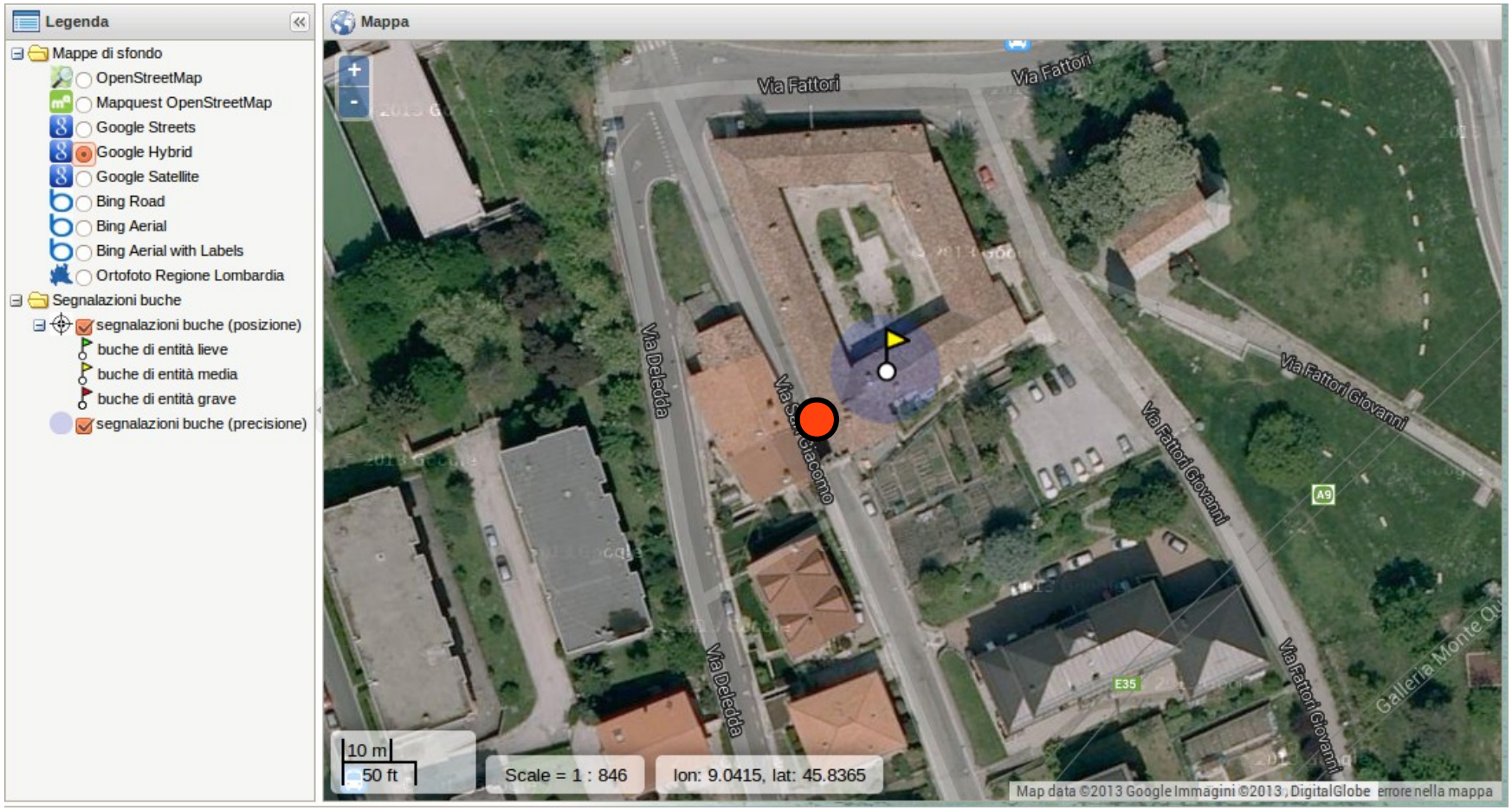
Results & considerations

- ✔ VGI data quality (positional accuracy) (Minghini 2014)
  - ➔ visualization of the accuracy circle, whose radius is given by Android



Results & considerations

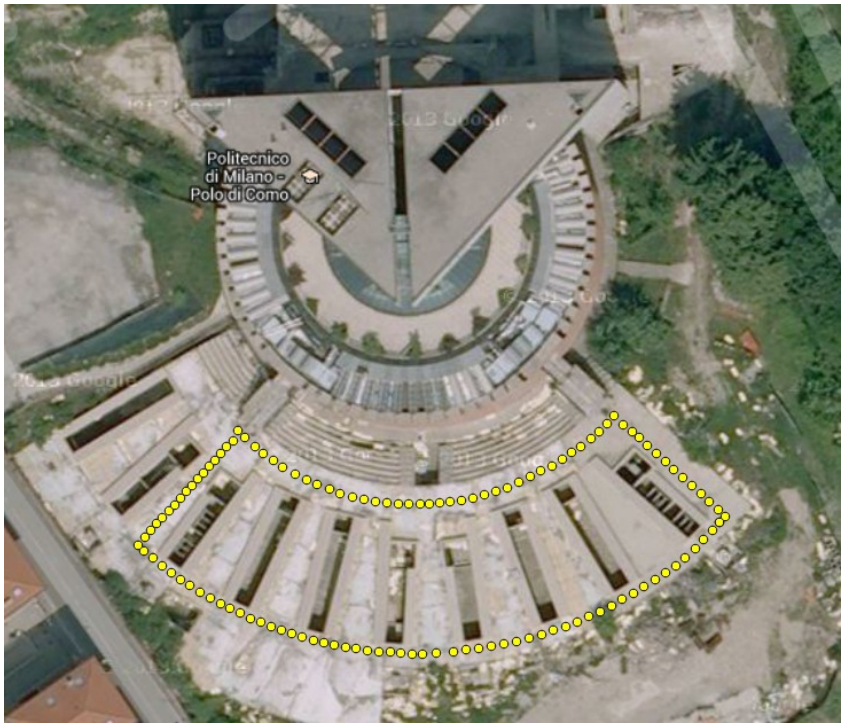
- ✔ VGI data quality (positional accuracy) (Minghini 2014)
  - ➔ visualization of the accuracy circle, whose radius is given by Android



  - ➔ *getAccuracy* returns the radius 68% confidence (one standard deviation if the sample is normally distributed) on horizontal position

Results & considerations

- GPS campaigns to compare the positions measured by
  - a high-accuracy GPS receiver: Leica RX1200 (positions taken as the truth)
  - a (low-accuracy) tablet: ASUS Eee Pad – GPS mode
  - a (low-accuracy) smartphone: Samsung Galaxy S Advance – A-GPS mode
- 3 repetitions with each mobile device on 2 sets of points



- Area 1 (122 points)

- Area 2 (62 points)

Results & considerations

- Main operations performed
  - outlier detection and removal through robust regression
  - ordinary linear regression between the high- and low-accuracy coordinates
  - test on the parameters of the ordinary linear regression
  - tests of independence/correlation between the X and Y errors of the mobile device estimated locations
- Main results
  - shorter TTFF (Time To First Fix) for A-GPS geolocation
  - variable number of outliers (tablet: even > 10%, smartphone: < 7%)
  - positional accuracy: median values of 4-10 m, max values of tens of meters
  - ordinary linear regression: significant disturbs due to both the position (*a* parameter) and a non-unique bias *b* lower than the positional accuracy
  - some correlation between the X and Y errors only for Area 2
  - variable reliability of the Android-returned accuracy value (tablet: even 30%, smartphone: > 75%)

Results & considerations

- Main operations performed
  - outlier detection and removal through robust regression
  - ordinary linear regression between the high- and low-accuracy coordinates
  - test on the parameters of the ordinary linear regression
  - tests of independence/correlation between the X and Y errors of the mobile device estimated locations
- Main results
  - shorter TTFF (Time To First Fix) for A-GPS geolocation
  - variable number of outliers (tablet: even > 10%, smartphone: < 7%)
  - positional accuracy: median values of 4-10 m, max values of tens of meters
  - ordinary linear regression: significant disturbs due to both the position (*a* parameter) and a non-unique bias *b* lower than the positional accuracy
  - some correlation between the X and Y errors only for Area 2
  - variable reliability of the Android-returned accuracy value (tablet: even 30%, smartphone: > 75%)