XV International Colloquium

The Role and Impact of International Institutions on Economic Theory and Policy

Maison des Sciences de l'Homme in Paris Nord, Paris – May 2-4, 2018

Social Power and Crowdsourcing

Rodolfo A. Fiorini

Social Power and Crowdsourcing

Virtual Presentation Topics

1. The Digital Currencies

2. The Fourth Industrial Revolutions

3. The Crowdsourcing Revolutions

4. The Information Copy (IC) Medicine Revolution

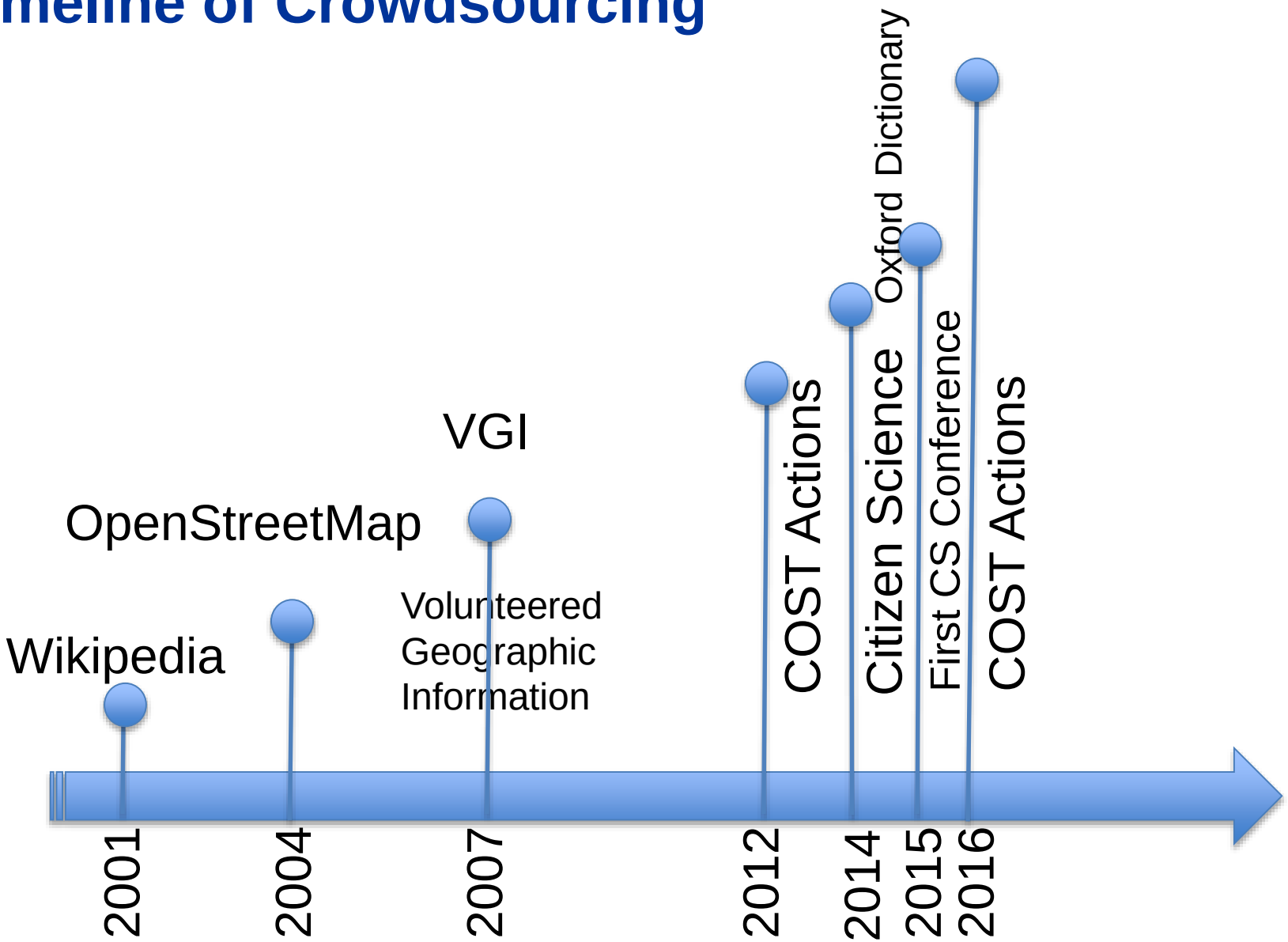
5. The Digitalisation of Public Services

6. The Incoming Super Smart Societies and Many More Promised Digitalisations…

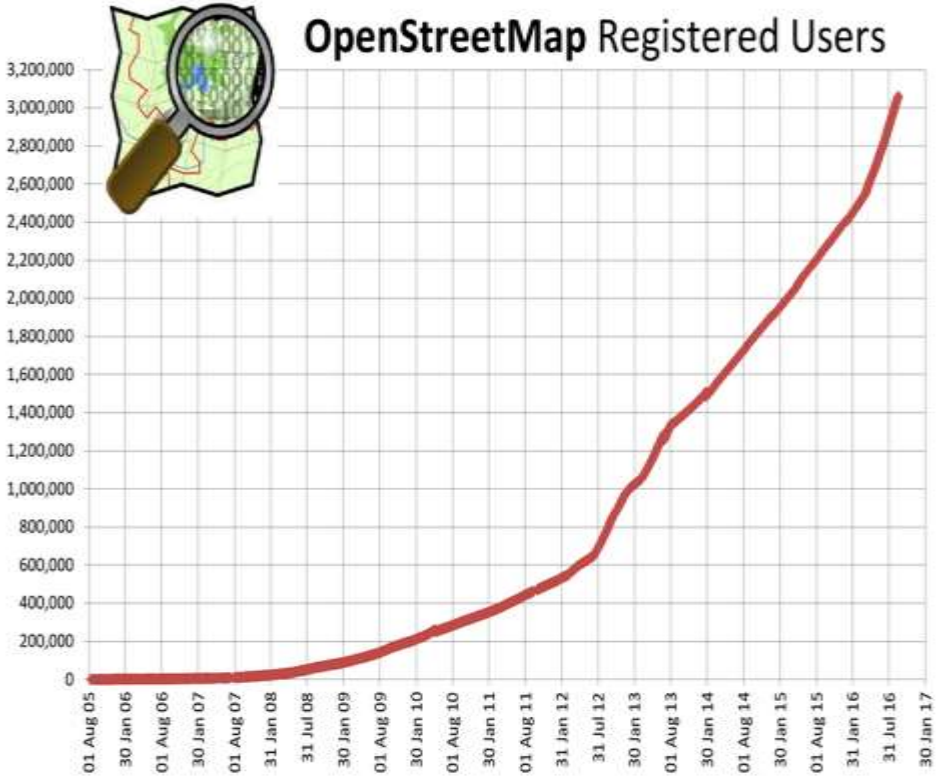
Presentation Outline

- ✓ Timeline of Crowdsourcing
- ✓ Crowdsourcing Enabling Factors
- ✓ Research Networks
- ✓ Citizen Science Projects
- ✓ Definitions and classifications
- ✓ Examples
- ✓ Open Questions
- ✓ Conclusions

Timeline of Crowdsourcing



OpenStreetMap Project



2001 2004 2007 2012 2014 2015 2016

Geospatial Information Crowdsourcing

- ✓ OpenStreetMap (OSM) is the most popular project of Volunteered Geographic Information (VGI). Born in 2004 for streets, then evolved into the largest, most diverse, most complete & most up-to-date geospatial database of the world.
- ✓ increasing number of contributors (currently over 3M)
- ✓ database available under an open license (ODbL)
- ✓ increasing interest from the academic community

Geo Crowdsourcing

- ✓ Geolocated sensors and handheld sensors: GPS, WiFi or Bluetooth receivers, cameras, microphones, activity trackers, sensor for body temperature, heart rate, brain activity, muscle motion and other critical data

Geospatial Information Crowdsourcing

- ✓ Geolocated sensors and handheld sensors: GPS, WiFi or Bluetooth receivers, cameras, microphones, activity trackers, sensor for body temperature, heart rate, brain activity, muscle motion and other critical data

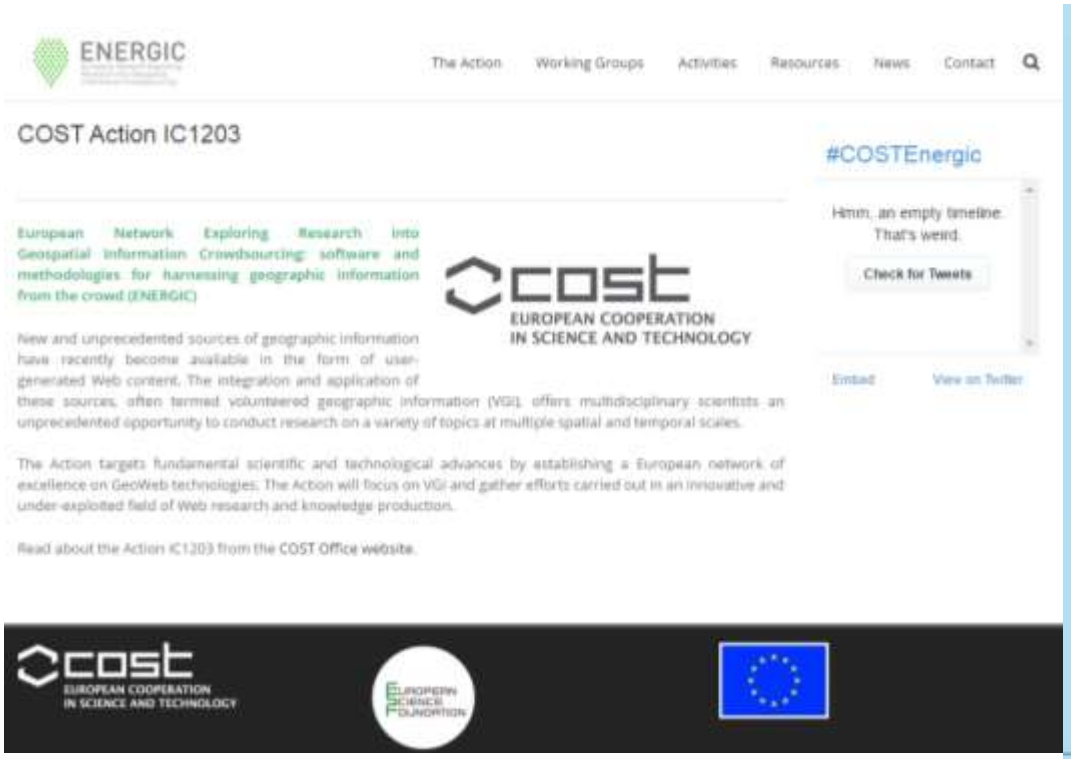
- ✓ Web2.0 technologies (Tim O'Reilly & Dale Dougherty 2004)

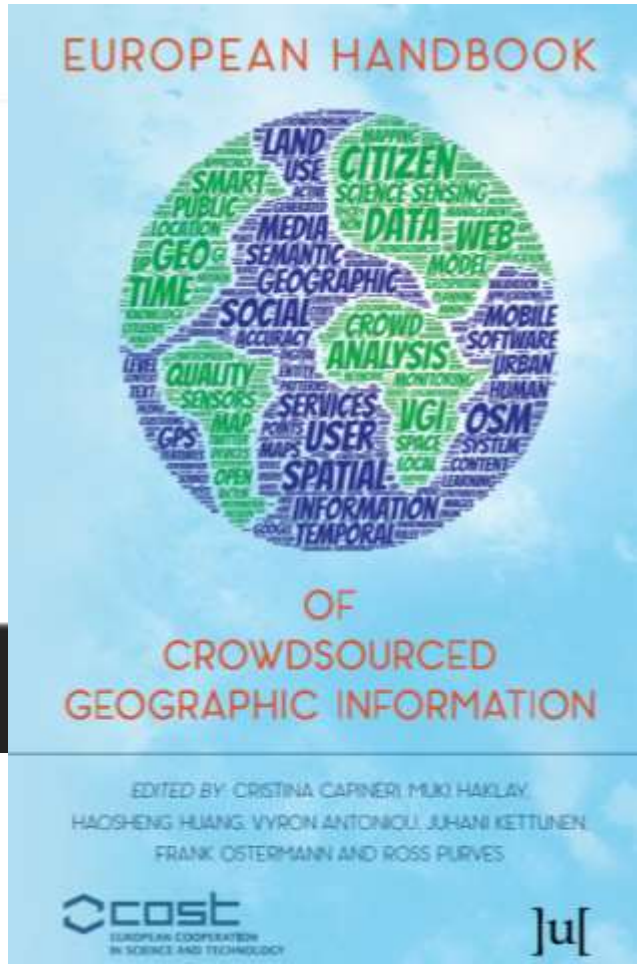
Geospatial Information Crowdsourcing

- ✓ Geolocated sensors and handheld sensors: GPS, WiFi or Bluetooth receivers, cameras, microphones, activity trackers, sensor for body temperature, heart rate, brain activity, muscle motion and other critical data
- ✓ Web2.0 technologies (Tim O'Reilly & Dale Dougherty 2004)
- ✓ Open access/collaborative and sharing approach to information resources: Collective Intelligence (Pierre Levy 1994)
- ✓ Sharing of skills, goods and services driven by the increasing sense of urgency of resource depletion

Pierre Levy, 1994, L'Intelligence collective. Pour une anthropologie du cyberespace, La Découverte, Paris.

Geospatial Information Crowdsourcing





Chair of the Action:
Prof Cristina CAPINERI (IT)

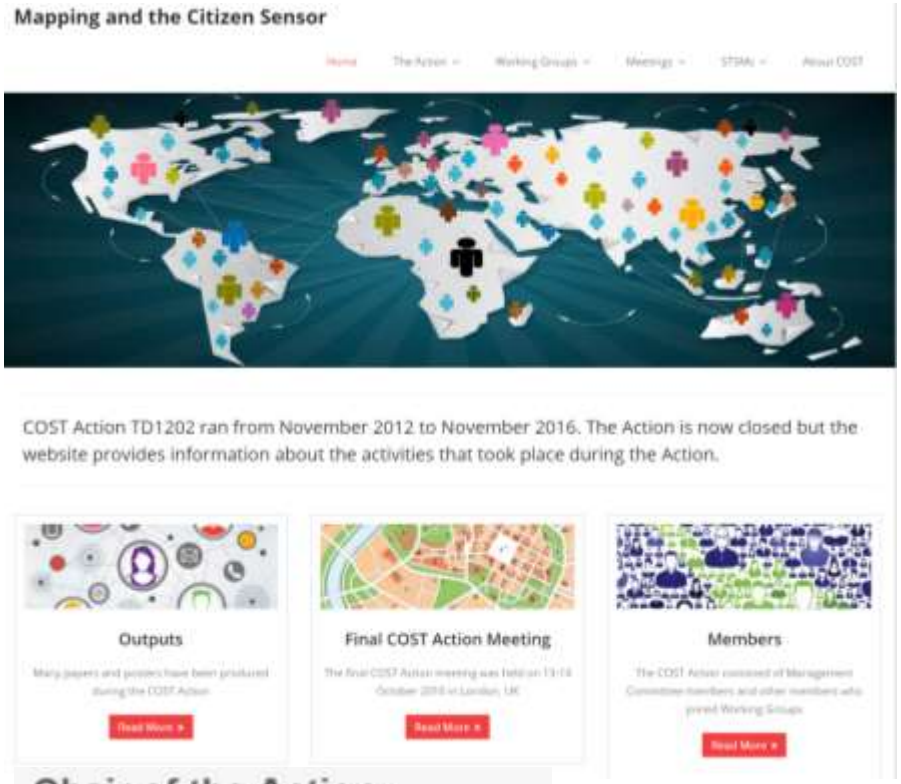
Vice Chair of the Action:
Prof Muki HAKLAY (UK)

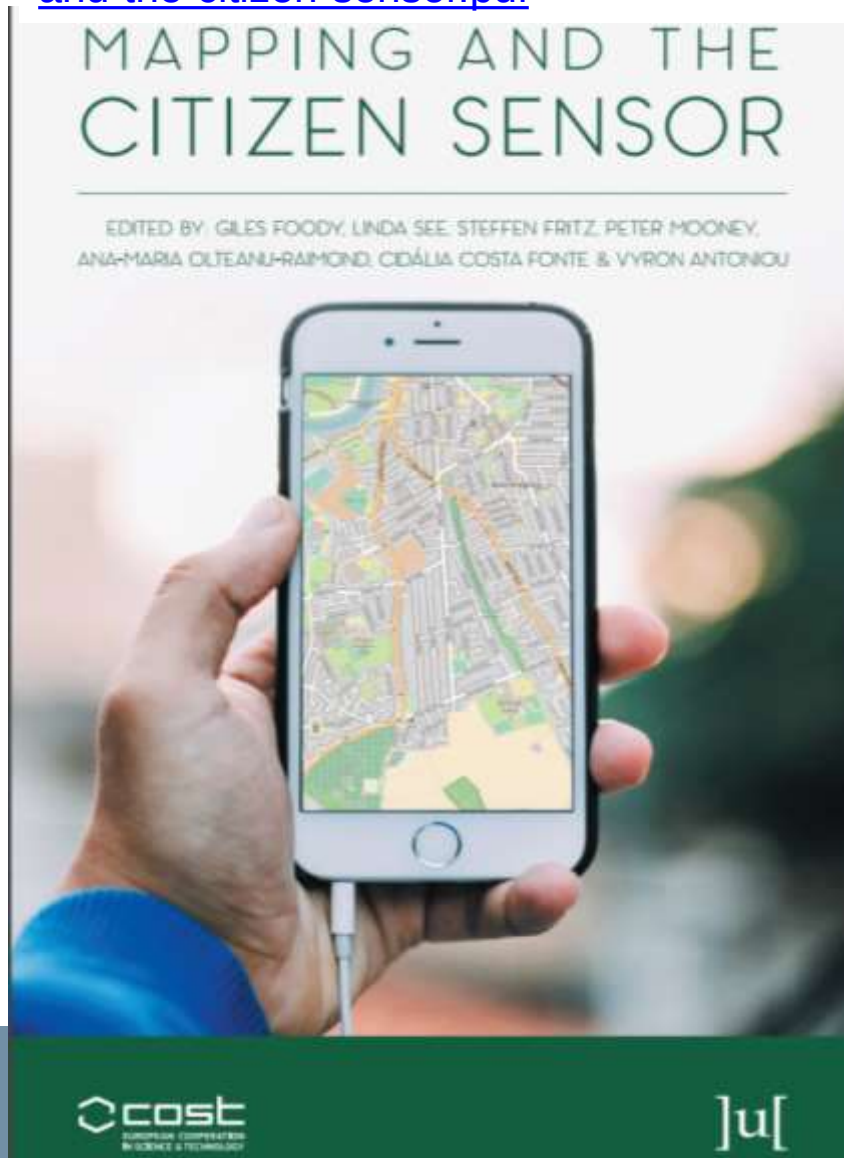
http://vgibox.eu/

http://www.ubiquitypress.com/site/books/10.5334/bax/

Geospatial Information Crowdsourcing

http://www.citizensensor-cost.eu/
http://pure.iiasa.ac.at/14814/1/mapping-and-the-citizen-sensor.pdf





The Citizen Science Cost Action

1. Ensure scientific quality of Citizen Science
2. Develop synergies with education
3. Improve society-science-policy interface
4. Enhance the role of CS for civil society
5. Improve data standardization and interoperability
6. Overarching - Cross-WG-Synthesis and overarching measures

Chair of the Action:
Dr Katrin VOHLAND (DE)

Vice Chair of the Action:
Dr Marisa PONTI (SE)

https://www.cs-eu.net/

The Citizen Science Cost Action

- ✓ A laundry list of projects available here: https://en.wikipedia.org/wiki/List_of_citizen_science_projects.
- ✓ Other sources:
  - ➔ Scistarter (https://scistarter.com/)
  - ➔ Citizen Science Alliance (https://www.citizensciencealliance.org/)
  - ➔ VGI Knowledge Portal (http://vgibox.eu/repository/index.php/Main_Page).

The Citizen Science Cost Action

✓**Citizen science**: set of practices in which citizens participate in data collection, analysis and dissemination of a scientific project (Cohn 2008)

✓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

Geospatial Information Crowdsourcing

- ✓ **Crowdsourced geographic information** → any data contributed by the crowd with a geographical reference (they could potentially be mapped)
  - → The geographic reference can be explicit or implicit (→gazetteer services, like GeoNames or Wikimapia).
- ✓ Data can be **actively** contributed (Volunteer Geographic Information) or **passively**.
- ✓ Data can be distinguished in:
  - → **Framework** data (those previously collected by NMAs: topographic databases, transportation networks, building footprints,etc)
  - → **No Framework** data (biodiversity, air quality, etc)

Geospatial Information Crowdsourcing

active

OpenStreetMap

CADASTA

OpenDroneMap

wikimapia

LAND COVER VALIDATION GAME

I CAMMINI DELLA REGINA

GeoNames

bikemap

GEO-Wiki

Citizen Weather Observer Program APRSWXNET

Map Share™ Reporter

TOMTOM

BIKE SHARE MAP

Ingress

Google Traffic

Google

tripadvisor

Panoramio

flickr

Instagram

TOMTOM

Pokemon go

Internet Live Statistics

- http://www.internetlivestats.com/
- 9,890 Tweets sent in 1 second
- 2,528 Instagram photos uploaded in 1 second
- 2,153 Tumblr posts in 1 second
- 1,843 Skype calls in 1 second
- 29,290 GB of Internet traffic in 1 second
- 50,232 Google searches in 1 second
- 106,299 YouTube videos viewed in 1 second
- 2,420,172 Emails sent in 1 second

Crowdsourcing Enabling Factors

- ✓ Providing something back as a motivator. This “something” has to be perceived as of greater value than the sustained effort.
  - → Personal satisfaction → make it as simple as possible to contribute
  - → Fun→ gaming
  - → Reputation → show immediately and to the whole community the individual contributions
  - → Monetary return → awards
  - → Gaining new knowledge
- ✓ Recruitment: it depens on the target audience.
  - → A launch event or side event at existing conferences, workshops and festivals
  - → Parties, like mapping parties or mapathons (armchair mapping parties)
  - → Schools and education

Mapathons and Humanitarian Mapathons

A **Mapathon** is a coordinated mapping event held generally **INDOOR**.

Humanitarian OpenStreetMap Team

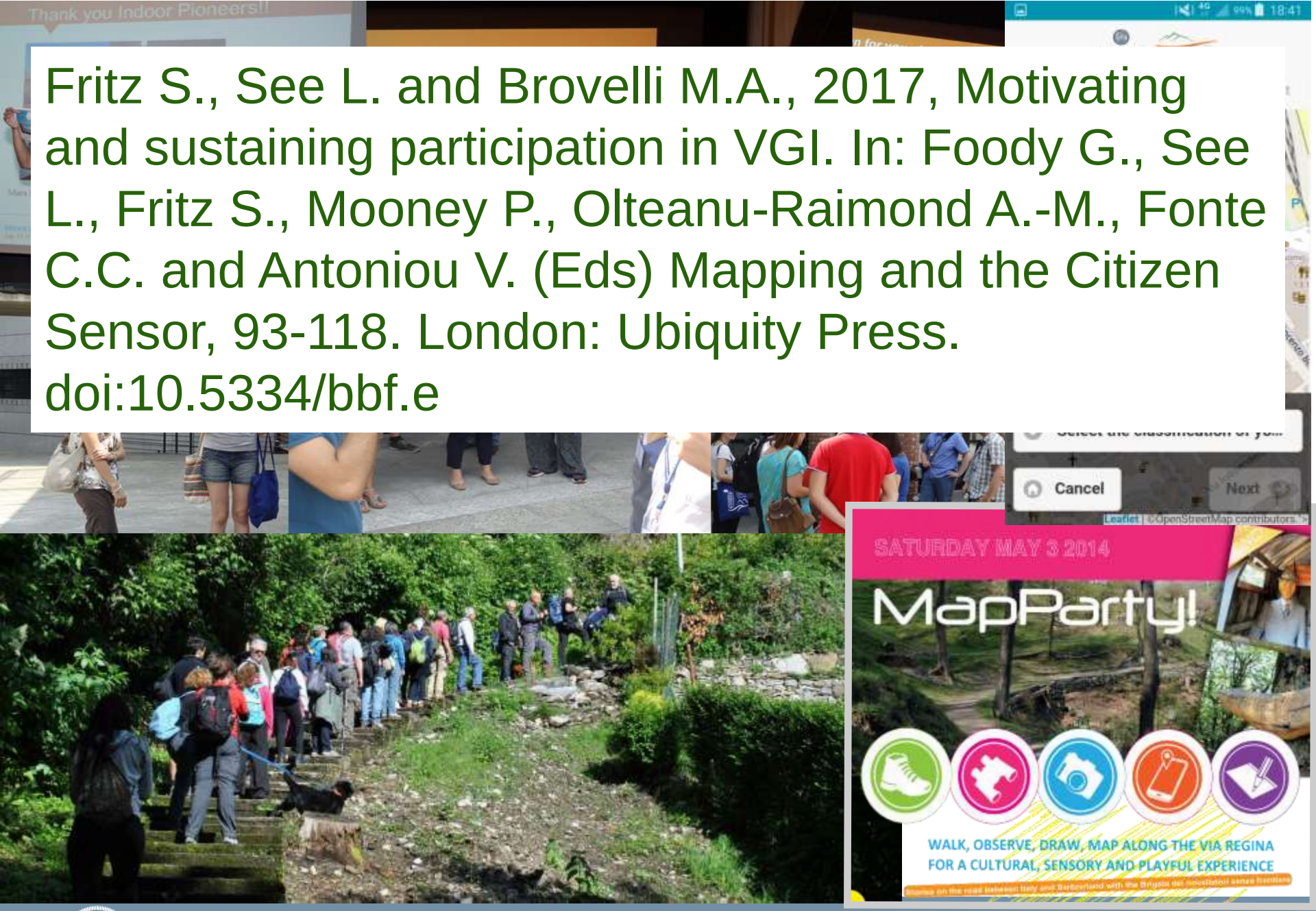
MISSING MAPS

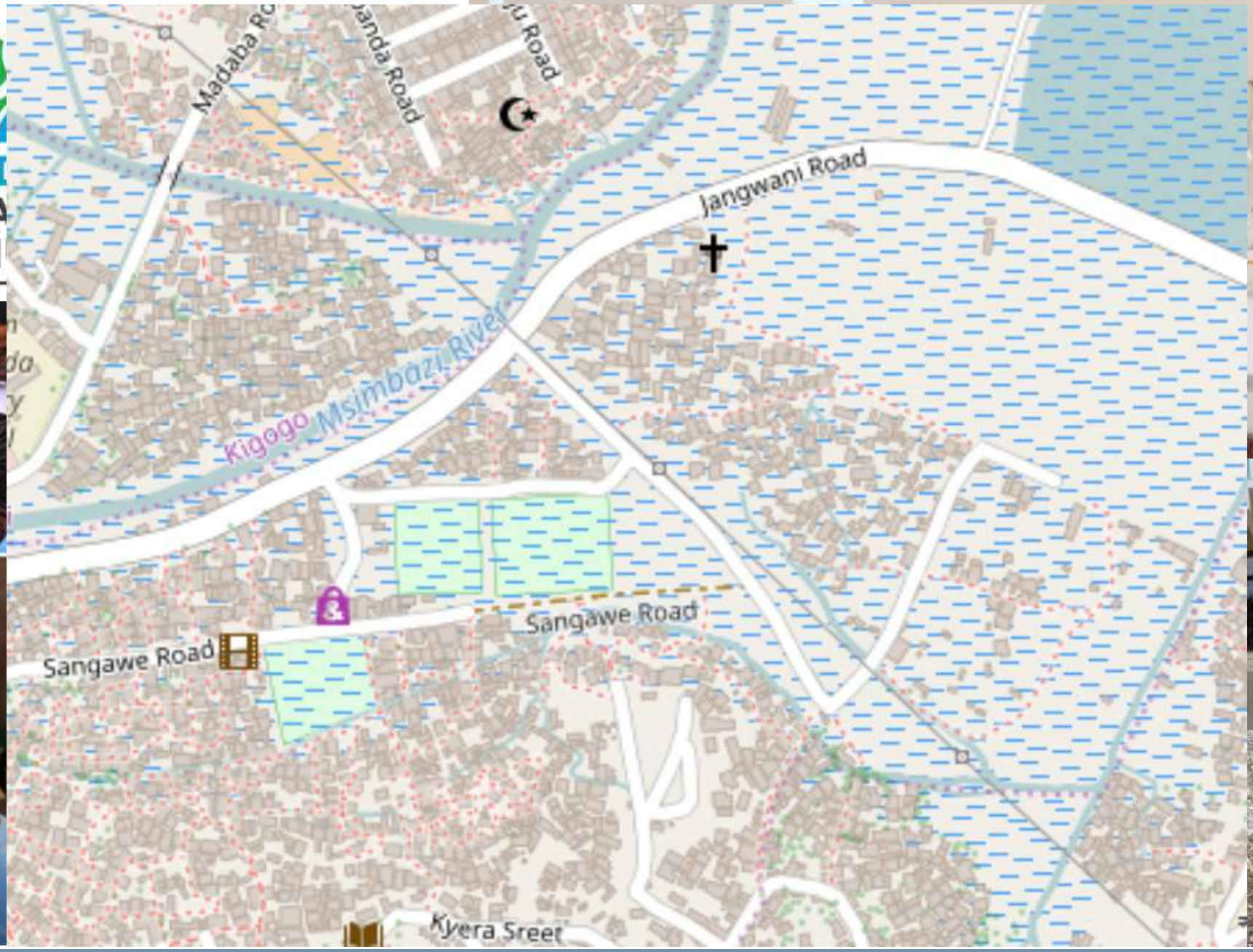
Step 1
Remote volunteers trace satellite imagery into OpenStreetMap

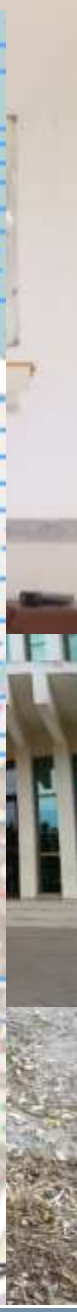
Step 2
Community volunteers add local detail such as neighborhoods, street names and evacuation centers

Step 3
Humanitarian organisations use mapped information to plan risk reduction and disaster activities that save lives.

Fritz S., See L. and Brovelli M.A., 2017, Motivating and sustaining participation in VGI. In: Foody G., See L., Fritz S., Mooney P., Olteanu-Raimond A.-M., Fonte C.C. and Antoniou V. (Eds) Mapping and the Citizen Sensor, 93-118. London: Ubiquity Press. doi:10.5334/bbf.e



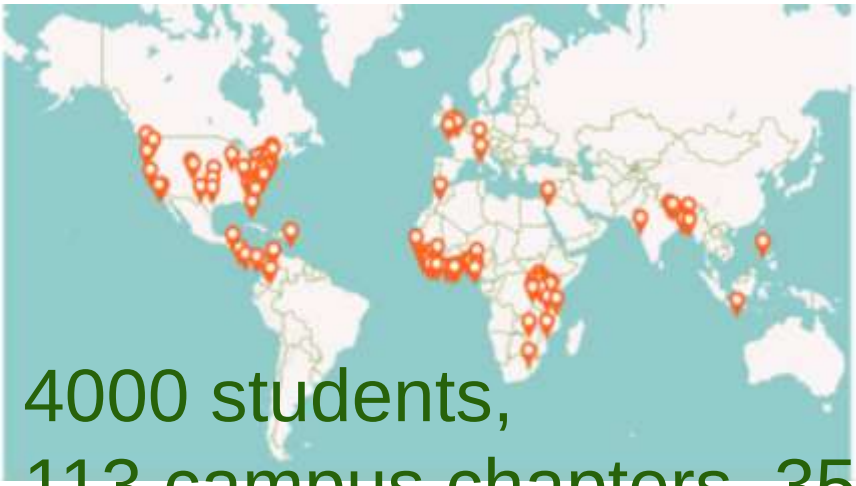

Madaba Road
Jangwani Road
Kigogo Msimbazi River
Sangawe Road
Sangawe Road
Kyera Sreet



Research Networks: YouthMappers

The motto of YouthMappers: WE DON'T JUST BUILD MAPS.
WE BUILD MAPPERS.

Capitalizing on web-based open geospatial technologies, the mission is to cultivate a generation of young leaders to create resilient communities and to define their world by mapping it.



4000 students,
113 campus chapters, 35 countries,

youthmappers

POLI MAPPERS

Research Networks: Minimapathons

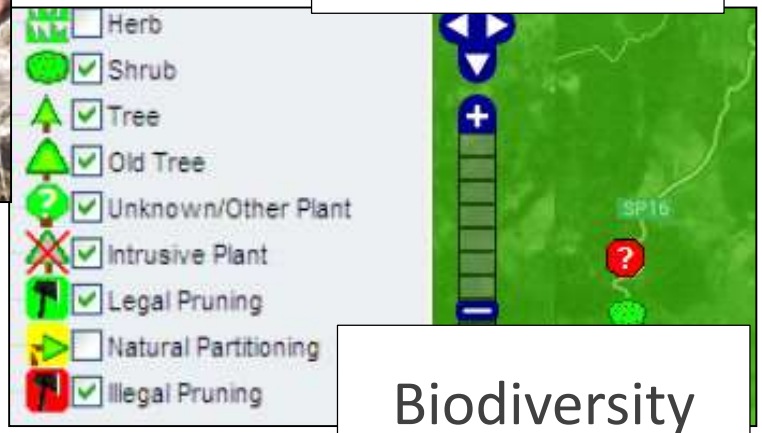
More than 200 kids mapping buildings in the northernmost part of Swaziland in a project for malaria elimination (task #1577)

Research Networks: Participatory Sensing

Architectural barriers

Cultural elements

Street furniture



Biodiversity

Research Networks: Open Street Map Geochicas

- OSM Latin American Women community **Geochicas** works towards closing the gender gap within the OSM community through different projects focused on the understanding of the role, participation and representation of women in OpenStreetMap.
- The **Geochicas working group** states that the data gathered by the OSM community is male oriented and biased given the fact that only 3% of OpenStreetMap contributors is perceived as woman and this has a direct impact on the information and points of interest that are mapped and how the map, the community activities are conducted

Crowdsourcing in National Mapping 2017

Held in Leuven, Belgium April 3rd and 4th 2017
With NMCAs, Geomatics Industry, academic research, software developers, citizens involved in geographic crowdsourcing and VGI, leaders or managers of crowdsourcing or VGI projects
http://www.cs.nuim.ie/~pmooney/eurosdr2017/

EuroSDR

The ESA Earth Open Science

NAVIGATION

- Home
- Objectives
- Committee
- Format
- Why attend?
- Sessions
- Abstract Submission
- Exhibition
- Contact

The ESA Earth Observation Φ-week EO Open Science and FutureEO > Sessions

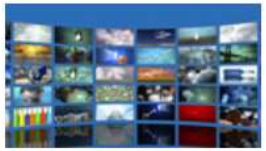
Open Science (OS)

The event includes community session by leading scientists on the following topics:

- **OS1** - AI & Data Analytics in and for Science: Exploring the transformational effects of AI and Data Analytics in EO research and Earth system science.
- **OS2** - Exploitation Platforms for collaborative Science: Addressing the advancements and opportunities offered by new Exploitation Platforms to advance EO research and promote collaborative approaches in Earth System Science.
- **OS3** - Data Cube: Exploring new approaches for Big Data processing exploiting latest advancements in Data Cube technologies.
- **OS4** - Citizen Science: Overview of the major citizen science initiatives, new opportunities for engaging citizens in science and EO research.
- **OS5** - Open Data, Tools and Virtual Labs: Exploring latest advancements in collaborative science offered by the combination of open data, new Open Tools and Virtual Laboratories.
- **OS6** - New Education: Addressing the latest advancements and opportunities in education activities offered by the combination of EO and advancements in ICT.
- **OS7** - Visualization and Science Communication: Overview of the latest advancement in visualisation of EO data and the opportunities offered by new visualisation tools in communicating science to the general public.
- **OS8** - Lightning Talks presenting scientific results of Early Career scientists on the above topics.

FutureEO (FE)

Video Talk Gallery



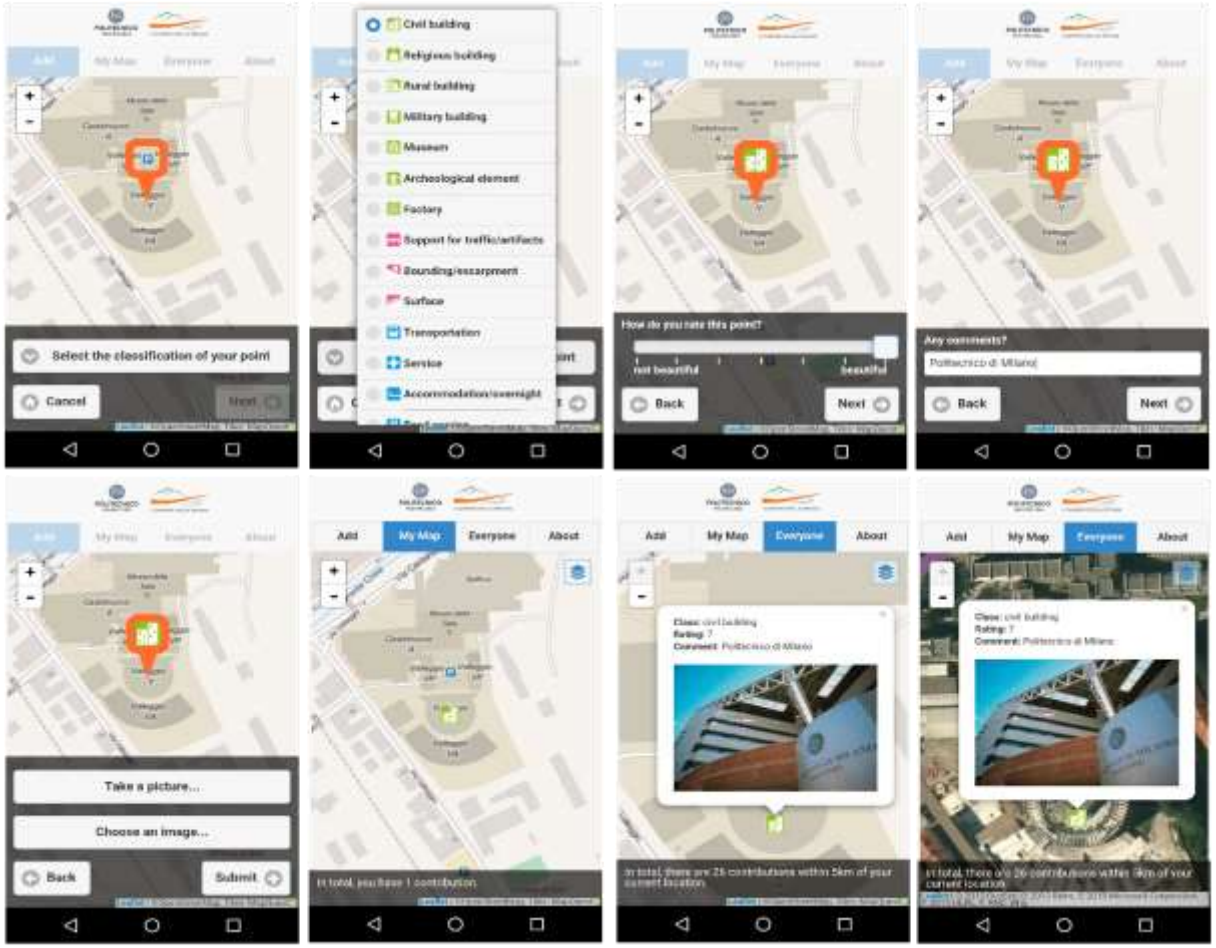
Video Poster Gallery

Social Media Story

EO OPEN SCIENCE

Examples

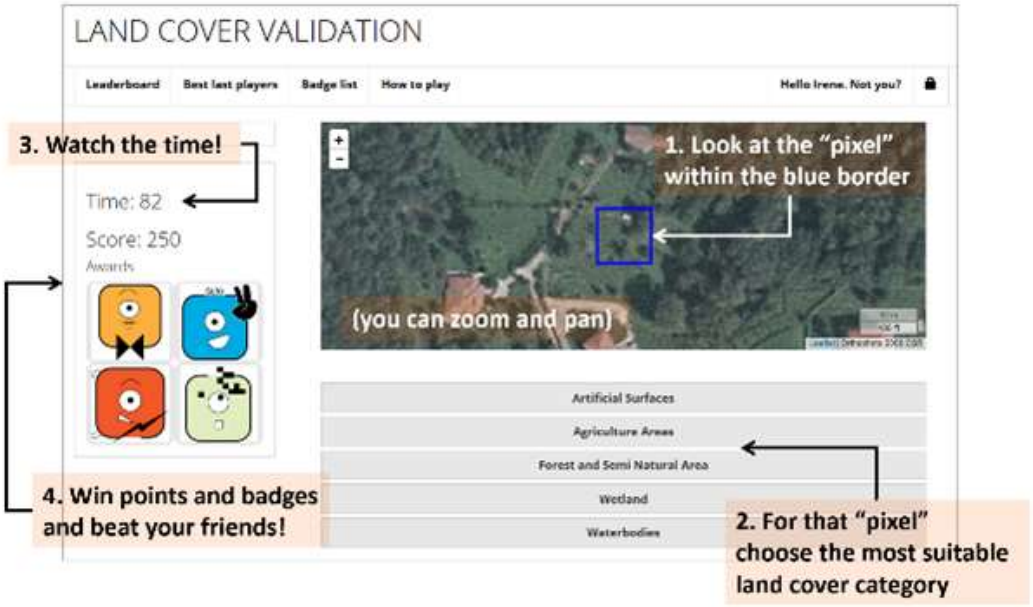
WebApp: http://viaregina3.como.polimi.it/app/ or App (Via Regina in Play and Apple Store)



Examples

http://131.175.143.84/land-cover-collector/

Examples



http://bit.ly/foss4game

LAND COVER VALIDATION GAME

MIGRATE

MIGRation pATterns in Europe

http://geomobile.como.polimi.it/migrate/

http://muvias.eoapps.eu/cityfocus/

CITY FOCUS

Ethical Issues

- ✓ It is an open question.

- ✓ The key ways to respect ethics in data-based research include:

  - ➔ involving participants throughout the research process;
  - ➔ avoiding collecting information that should remain private, notifying participants of their inclusion and providing them with options to correct or delete personal information
  - ➔ using public channels to disseminate research such as Open Data.

Privacy

- ✓ Private data is any data or information that can be linked to an individual contributor who created, collected or edited that data.
- ✓ Many citizens (even when active contributors) are not aware about the possible downstream future usages of their content.

Graffiti in Shoreditch, London – Zabou (Wikimedia Commons)

- ✓ We are constantly tracked by the devices that we carry around with us. Is location information, in itself, a private data or can it be linked to individuals? That depends on the location accuracy.
- ✓ How to protect not only the privacy of the volunteer contributing citizen but also the privacy and security of the subjects (human and non-human) of the contribution?

Privacy-Friendly Apps and Protocols

To avoid compromising privacy and security, there is an urgent need for privacy-friendly applications and protocols.

Technological approaches

- ✓ Blurring or fuzzing;
- ✓ Anonymizing data and selectively revealing information according to citizen preference;
- ✓ Privacy-preserving data mining techniques:
  - ➔ appropriate levels of anonymity , to avoid the undesired side effects on privacy, by means of controlled transformation of data and/or patterns
  - ➔ but with limited distortion to preserve the possibility of discovering useful patterns and trends.

Legal Issues: Data Property

- Share alike licenses that require the derived datasets to be released with the same license; the most famous one in geographical information is the Open Database License (ODbL) used by OpenStreetMap.

As long as you:

Attribute: You must attribute any public use of the database, or works produced from the database, in the manner specified in the ODbL. For any use or redistribution of the database, or works produced from it, you must make clear to others the license of the database and keep intact any notices on the original database.

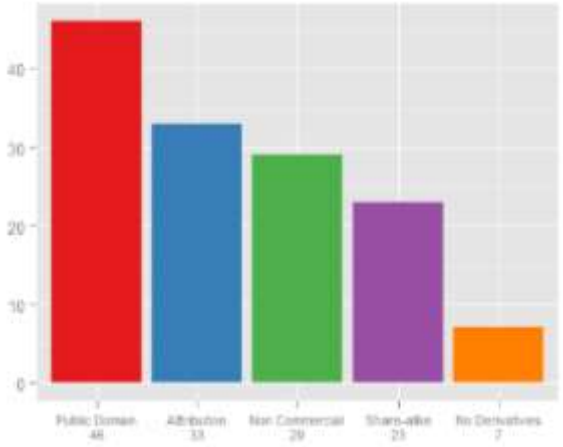
Share-Alike: If you publicly use any adapted version of this database, or works produced from an adapted database, you must also offer that adapted database under the ODbL.

Keep open: If you redistribute the database, or an adapted version of it, then you may use technological measures that restrict the work (such as DRM) as long as you also redistribute a version without such measures.

Legal Issues: Data Property

Survey report: data management in Citizen Science projects – JRC (Sven Schade and Chrysi Tsinaraki, 2016)

- ✓ public domain, i.e. completely free from any restriction of intellectual property;
- ✓ with attribution, i.e. giving credit to the original creator;
- ✓ as share-alike, i.e. licensing derivatives under identical terms;
- ✓ non-commercial, i.e. allowing any re-use that is not of a commercial nature;
- ✓ no derivatives, i.e. preventing any way of changing the original source or building upon it.

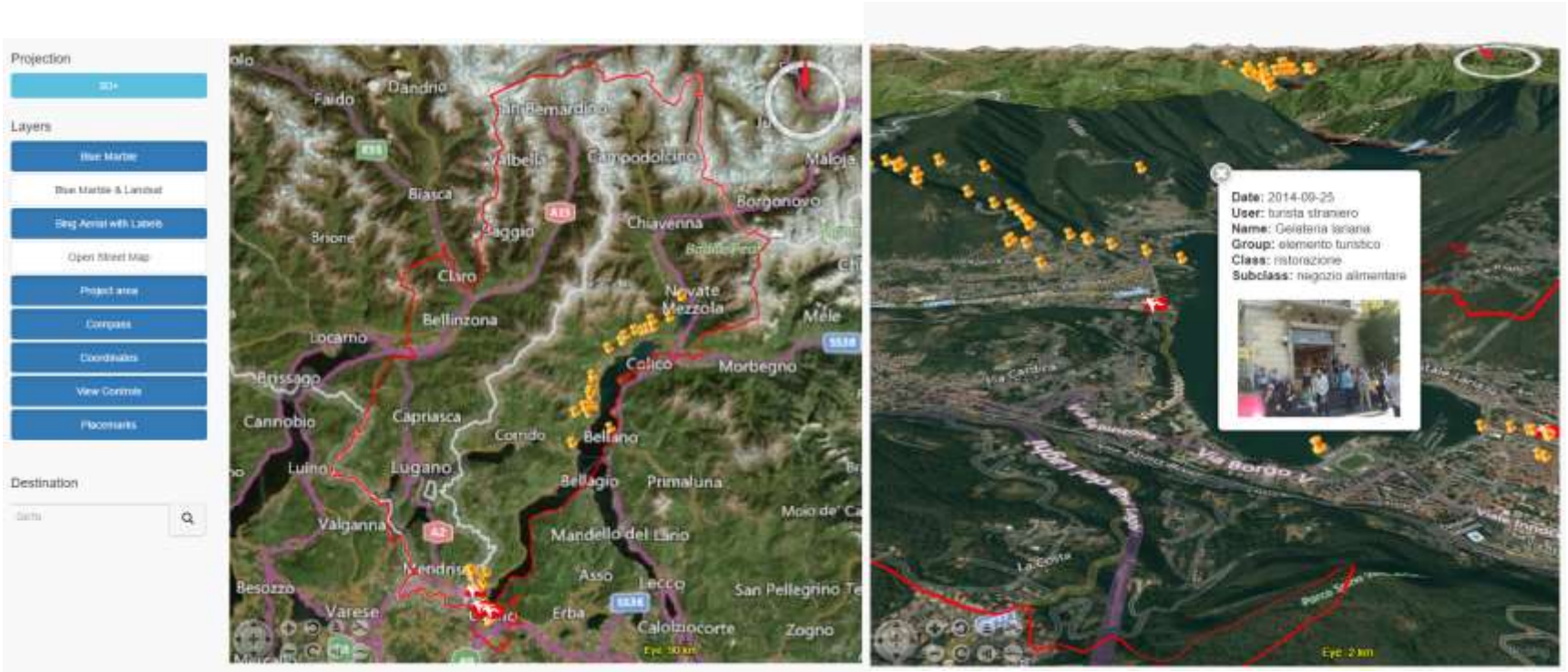


http://publications.jrc.ec.europa.eu/repository/handle/JRC101077

Legal Issues: Responsibility

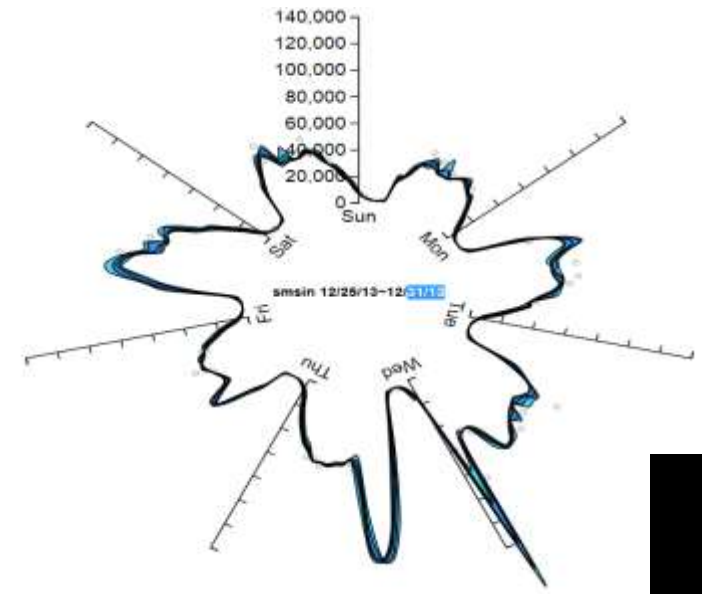
- ✓ The question of intellectual property is strictly related to the responsibility.
- ✓ Who is responsible for the correctness and reliability of data?
- ✓ Many citizens are not fully aware of the additional intelligence that can be elicited to they content by the powerful combinations of cloud computing and data mashuping and processing technologies available today.
  - ➔ From the position of the volunteer, their legal role and their contribution may not always be clearly defined and this can lead to potentially exposing them to legal problems.
  - ➔ On the other hand, if a data provider or data portal only facilitates the transfer or access to VGI data, then who carries the legal responsibilities related to consequences of future use of these data?

Visualisation



http://viaregina3.como.polimi.it/WorldWind/

Visualisation



Received SMS texts
from Wednesday, December 25, 2013
to Tuesday, December 31, 2013
for all Milano grid cells

Received Calls (6 days)
Color: # SMS
X,Y: lat,long
Z: day of the week

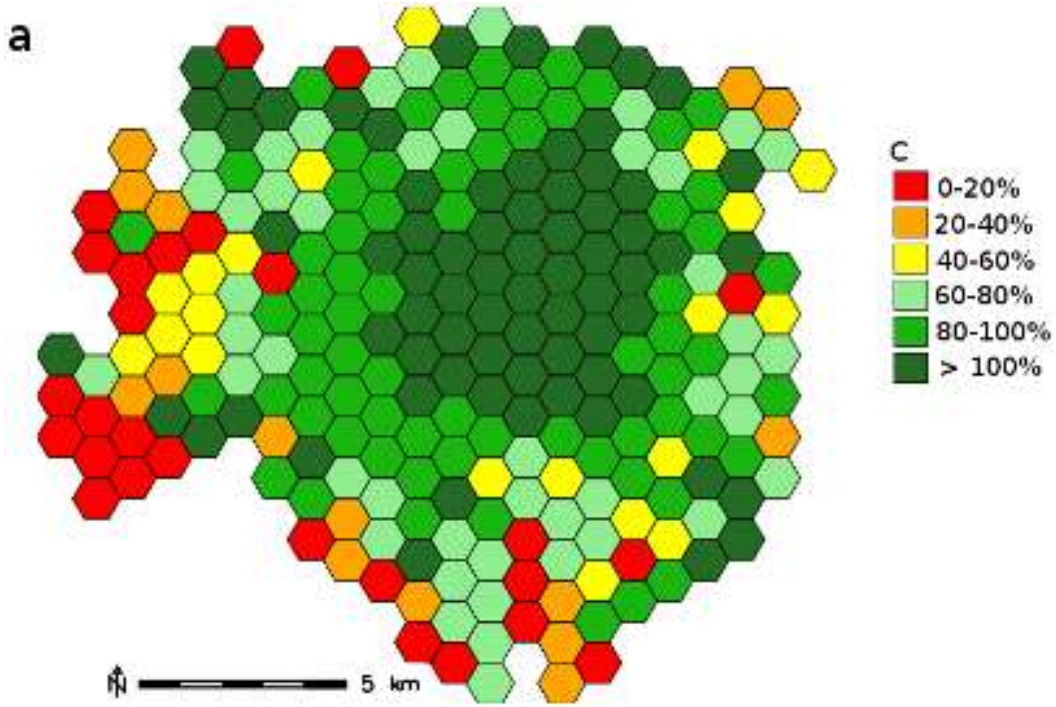
Quality

- ✓ ISO Principles and Guidelines.
- ✓ Reference Datasets.

- ✓ Fundamental Assumption: authoritative data are of higher quality than crowdsourced (is it true??)
- ✓ ISO Quality Elements:
  - ➔ Completeness
  - ➔ Logical Consistency
  - ➔ Thematic Accuracy
  - ➔ Positional Accuracy
  - ➔ Temporal Accuracy
  - ➔ Usability

Completeness

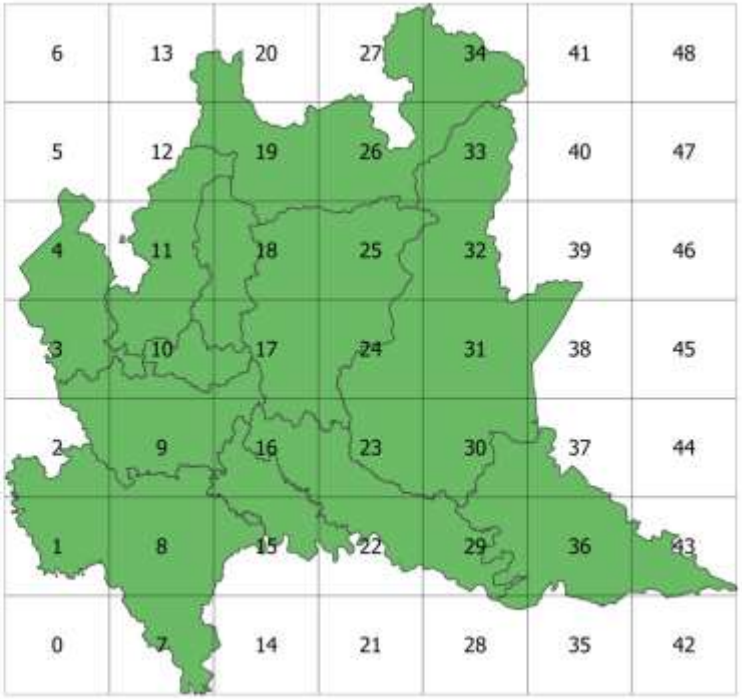
Completeness: OSM vs Authoritative Topographic DB at scale 1:2000, Milan 2016 (building footprints)



C VALUES	%
> 100%	28.9%
80% < C < 100%	27.7%
60% < C < 80%	18.5%
40% < C < 60%	8%
C < 40%	16.9%

Positional Accuracy

Positional Accuracy: mean and RMS of the distance between homologous pairs of OSM and the Authoritative Maps in case of building footprints



Building footprints in

- OSM
- Topographic Database DBT (scale: 1:1000→ 1:5000)

Area	DBT buildings	DBT points	OSM buildings	OSM points
23.844 Km^2	2794196	17604429	915696	7506851

TILEs	# Buildings OSM= DBT	# Buildings OSM	%
1:	0	6.807	0,0%
2:	21.902	31.131	70,4%
3:	11.667	98.130	11,9%
4:	531	40.717	1,3%
5:	0	381	0,0%
7:	1	36.072	0,0%
8:	5	49.345	0,0%
9:	1.298	71.402	1,8%
10:	15.689	168.830	9,3%
11:	13.310	49.383	27,0%
12:	295	11.014	2,7%
13:	0	1.730	0,0%
14:	0	932	0,0%
15:	3.940	8.487	46,4%
16:	11.995	55.334	21,7%
17:	4.982	61.569	8,1%
18:	13.567	26.736	50,7%
19:	2.047	23.701	8,6%
20:	0	1.190	0,0%

TILE	# Buildings OSM= DBT	# Buildings OSM	%
22:	5.316	10.709	49,6%
23:	0	8.696	0,0%
24:	671	34.033	2,0%
25:	0	8.301	0,0%
26:	352	14.630	2,4%
27:	0	1.858	0,0%
28:	0	625	0,0%
29:	0	2.281	0,0%
30:	7.325	23.003	31,8%
31:	1.364	33.616	4,1%
32:	0	8.950	0,0%
33:	0	4.817	0,0%
34:	0	3.320	0,0%
35:	2.247	2.415	93,0%
36:	4.826	15.607	30,9%
37:	0	6.414	0,0%
38:	0	4.322	0,0%
39:	0	1.200	0,0%
43:	10.176	11.847	85,9%

TILE	Points	m(DX)	var(DX)	m(DY)	var(DY)	m(dist)	var(dist)	min(dist)	max(dist)	PtiDBT%)	PtiOSM(%)
1	2359	0,588	2,070	-0,151	1,644	1,669	1,296	0,014	7,340	84,6	44,7
2	33540	-0,107	1,549	-0,560	1,532	1,465	1,261	0,000	6,634	72,8	23,6
3	372232	-0,118	1,992	-0,645	1,599	1,680	1,200	0,000	6,841	83,5	48,9
4	152706	0,015	2,753	-0,817	2,414	2,076	1,526	0,000	7,159	82,3	37,2
5	1366	-0,544	1,988	-0,547	1,836	1,776	1,264	0,022	6,673	80,5	35,9
7	117271	0,542	2,836	-0,344	2,168	1,999	1,420	0,000	7,361	75,1	62,6
8	170223	-0,013	1,670	-0,348	1,107	1,363	1,039	0,000	6,704	79,6	70,8
9	354194	0,235	1,585	0,024	1,689	1,430	1,285	0,000	6,701	79,5	44,9
10	568039	0,229	2,285	-0,375	2,318	1,770	1,664	0,000	6,798	77,6	40,0
11	128766	-0,019	2,740	-0,751	2,927	2,075	1,925	0,000	7,318	77,9	34,6
12	27452	0,556	4,112	-0,970	3,667	2,539	2,580	0,014	7,906	78,0	40,1
13	2394	-0,252	3,270	-0,898	1,583	2,054	1,503	0,010	6,780	87,8	47,9
15	18347	0,215	1,388	-0,671	1,697	1,486	1,374	0,000	6,621	85,6	40,1
16	179028	0,070	1,490	-0,803	2,420	1,706	1,651	0,000	7,469	83,4	53,3
17	216559	0,325	2,656	-0,689	2,364	1,979	1,682	0,000	7,089	83,4	40,9
18	43736	-0,060	2,871	-0,776	2,765	2,076	1,933	0,000	7,460	82,4	21,7
19	72066	0,319	3,183	-0,514	2,027	1,998	1,583	0,000	6,857	82,4	39,9
20	1684	0,763	3,136	-0,234	2,407	2,147	1,570	0,051	6,862	85,6	39,6

Total number of Points > 3 M

TILE	Points	m(DX)	var(DX)	m(DY)	var(DY)	m(dist)	var(dist)	min(dist)	max(dist)	PtiDBT%	PtiOSM(%)
22	20574	0,678	2,247	-0,214	1,654	1,597	1,854	0,000	7,312	78,6	42,4
23	34975	0,162	2,760	-0,780	3,646	2,216	2,129	0,000	7,784	78,7	50,1
24	133993	0,535	2,839	-1,254	3,324	2,436	2,085	0,000	7,641	81,5	42,3
25	17660	0,413	3,458	-0,739	3,309	2,395	1,747	0,010	7,567	83,1	36,2
26	52071	-0,281	3,291	-0,766	2,726	2,200	1,842	0,000	7,511	79,0	43,9
27	7840	0,559	4,841	-0,399	3,259	2,593	1,846	0,022	7,048	86,5	54,4
28	246	1,326	1,605	-0,228	0,980	1,836	1,019	0,054	7,112	80,4	51,7
29	5998	0,293	4,094	0,390	1,873	2,133	1,657	0,010	6,742	81,0	53,9
30	58863	0,406	3,169	-1,105	4,635	2,606	2,396	0,000	8,034	82,4	37,8
31	128898	0,578	3,696	-0,742	3,575	2,457	2,118	0,000	7,734	78,2	42,7
32	26561	0,427	4,010	0,270	2,987	2,343	1,761	0,014	8,081	81,4	50,1
33	15751	-0,253	2,943	-0,411	3,116	2,151	1,664	0,010	6,789	82,3	38,5
34	12746	-0,428	3,268	-0,460	3,291	2,304	1,645	0,014	7,562	85,0	38,5
35	109	-0,338	1,934	-0,114	1,044	1,342	1,287	0,000	5,237	58,3	0,9
36	21747	-0,214	2,473	0,131	2,078	1,761	1,513	0,000	6,427	71,0	35,5
37	6568	-0,330	3,765	0,961	2,611	2,381	1,739	0,022	8,133	79,1	47,9
38	11646	0,847	4,405	-0,037	2,632	2,413	1,931	0,000	9,533	76,6	40,3
39	2697	0,051	4,670	0,759	1,685	2,275	1,759	0,051	8,162	80,8	41,1
43	932	1,243	4,709	-0,128	1,405	2,227	2,713	0,000	8,623	59,1	1,8

Accuracy

	# BUILDINGS OSM = DBT	# BUILDINGS OSM	%
BOSTON	1.359	98.326	1,4%
S. FRANCISCO	9.722	166.764	5,8%

TILE	Points	m(DX)	var(DX)	m(DY)	var(DY)	m(dist)	var(dist)	min(dist)	max(dist)
Boston	446241	0,163	2,049	-0,162	1,953	1,681	1,228	0,000	6,422
SanFran	1E+06	-0,611	1,005	0,165	0,345	0,900	0,939	0,000	7,024

Open Geospatial Consortium Standards

OGC Citizen Science DWG will address the citizen science relevant aspects of interoperability:

- ✓ Hardware communication (standards used by sensors communicate e.g to a mobile phone);
- ✓ Data acquisition (how devices send data to repositories);
- ✓ Data storage and dissemination (how repositories make data discoverable and available);
- ✓ Data curation and preservation (how the data is maintained in particular in the long term when the actual data campaign is finished).
- ✓ http://external.opengeospatial.org/twiki_public/CitizenScienceDWG/WebHome

Conclusions

- ✓ Crowdsourcing is relatively new but it is a high-pace evolving approach to science and research.
- ✓ Many applications have been exiting the scientific field and are becoming of high interest in the business world.
- ✓ Despite the (many) projects, it is still a research in progress in various fields, from the legal and ethical point of view to the more technical questions.
- ✓ Being highly related to sensors (IoT), connectivity (the Cloud), volume, variety, velocity and veracity (Big Data) we are just at the very beginning and we expect a great evolution in next years.

Social Power and Crowdsourcing