Science 2.0 White Paper: Citizen Science (Session B.1)

Linda See[1], Steffen Fritz[1], Heather Leson[2], Stuart Wrigley[3], Brooke Simmons[4], Frans Snik[5], Thomas Bartoschek[6], Phillip Schneider[7], Alejandro Sánchez de Miguel[8] and Eleanor Bruce[9]

1 International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
2 Qatar Computing Research Center
3 University of Sheffield, UK
4 University of Oxford, UK
5 University of Leiden, The Netherlands
6 Institute for Geoinformatics, University of Münster, Germany
7 Norwegian Institute for Air Research (NILU), Oslo, Norway
8 Universidad Complutense de Madrid, Spain
9 University of Sydney, Australia

1 State-of-the-art in Citizen Science for EO Open Science

There were ten presentations in the citizen science session followed by a roundtable discussion. A summary is provided below, grouped by theme.

1.1 Citizen Science to Support Earth Observation (EO)

The Geo-Wiki suite of tools was presented by Steffen Fritz of the International Institute of Applied Systems Analysis (IIASA). Geo-Wiki is an application for visualization, crowdsourcing and validation of land cover originally developed for improving global land cover datasets. Since then a number of tools have been added, in particular the LACO-Wiki tool, which allows anyone to upload a land cover map, generate a sample, validate the map using very high resolution imagery and generate accuracy statistics, as well as the most recent serious game applications: Cropland Capture, FotoQuest Austria and Picture Pile. Launched during the Science 2.0 conference, Picture Pile is designed to be a generic application that involves the crowd in classification of imagery and geo-tagged photographs. The current focus is on asking citizens to help improve maps of deforestation but any type of images can be classified. We see Picture Pile as a potentially useful tool for ESA to validate many different products. There is also potential for Picture Pile to be valuable for humanitarian use, e.g. rapid damage assessment from UAVs and geotagged photographs.

Another presentation directly related to citizen science support to EO was given by Stuart Wrigley from the University of Sheffield. He presented the Crowd4Sat project, currently funded by ESA, where the objective of the project is to explore new ways of using citizen science in ESA's products and services, i.e. space data validation and space data exploitation as well as demonstrating the value of citizen science in scientific research and education. This objective will be achieved via four use cases: (i) snow coverage monitoring in the Spanish Pyrenees; (ii) traffic and pollution mapping and management; (iii) flood emergency mapping; and (iv) land use mapping for validation of CORINE land cover. These different use cases will make use of data and products derived from different satellites, e.g. Sentinel 1 and MODIS, help to improve the validation of products and services, and provide citizens with access to information, e.g. safe routes for hikers, more accurate flood risk maps, etc.

The iSPEX project was presented by Frans Snik, which involves citizens in the monitoring of aerosol optical thickness (AOT) using a low cost add-on to an iPhone. Although it is known that particulate matter has a negative effect on human health, more measurements are needed at a high spatial and temporal resolution in order to better understand this relationship. Other reasons include the need to better understand the effect of volcanic ash on jet engines and the effect of aerosols on climate change. Successful iSPEX campaigns have involved thousands of people in taking measurements across the Netherlands. When these measurements were compared to AOT from MODIS, there was general agreement in the patterns. Agreement between citizen measurements (once averaged across at least 50 people) and ground reference measuring stations was also extremely high. iSPEX campaigns have been rolled out to different cities across Europe, where thousands of measurements have been collected since the beginning of September.

The final presentation of direct relevance to EO is the Cities at Night project, presented by Alejandro Sánchez de Miguel. Using an off-the-shelf crowdsourcing platform to build the application, citizens help classify and georeference photographs from the International Space Station (ISS) that are taken at night. These photographs have a higher resolution than satellite-based nighttime light products as well as true colors. These photographs have already been used in a wide range of applications from helping municipalities to save energy, research on breast cancer, reducing air pollution, mitigating light pollution and increasing road safety to name but a few. There is currently a lack of funding to continue the project despite its considerable potential.

1.2 Citizen Science to Support Disasters

The keynote by Heather Leson focused on human computing in disaster assessment, with the aim of training algorithms to do these tasks automatically in the future. She showed the Micromappers tool, where citizens examine tweets and determine if they are important for humanitarian response. In this way the information is curated by the community and only the most important information, in particular where critical damage has been identified, is passed onto responders. These data are then used to train a classifier so that it automatically learns to identify actionable emergency information. A second example uses citizens to digitize UAV imagery for damage in order to help train computer vision algorithms. UAV imagery is now being shared openly via the Open Aerial Map initiative from Humanitarian OpenStreetMap (HOT). An example of how citizens working remotely were trained by local staff to provide maps after the Nepal earthquake (via the Missing Maps initiative) showed the effectiveness of citizen involvement around the world after this disaster. Heather asked how we can involve more people in analyzing satellite imagery. She suggested holding satellite image camps where people are trained to work with satellite imagery, in particular in low bandwidth areas. Heather's key message was that the EO community needs to respond with the next generation tools and training materials. Furthermore, she stressed HOT's need for processing tools and a stronger collaboration with the research community.

Brooke Simmons of Zooniverse/Oxford University presented examples of new humanitarian Zooniverse applications. Working with Humanitarian OpenStreetMap (HOT), the Zooniverse volunteers helped to create maps of population density, which showed areas of potential vulnerability that had not yet been mapped by HOT in detail. This information provided responders with locations of where water supplies should be sent. Another application provides citizens with UAV imagery, degraded to the resolution provided by Planet Labs, in order to examine how well citizens can identify damage at different resolutions and to train computer vision algorithms to identify damage.

Finally, Eleanor Bruce of the University of Sydney has been involved in the development of a community-based project for bushfire preparation. Rather than focusing on post-disaster, this citizen science project aimed to build resilience to potential bushfires and prepare citizens for events before they happen.

1.3 Citizen Science to Support Environmental Monitoring

A number of presentations were focused on how citizens can be involved in environmental monitoring. Two presentations were from ongoing FP7-funded citizen observatories. Stuart Wrigley of Sheffield University presented the tools developed in the WeSenseIt citizen observatory while Phillip Schneider of the Norwegian institute of Air Research presented work from the Citi-sense citizen observatory. The WeSenseIt citizen observatory is built around a two-way communication paradigm between citizens and authorities in order to better manage water resources. One aspect of the project has focused on driving down the cost of sensors while the other has concentrated on harnessing the collective intelligence of citizens through participatory sensing, activity understanding and large scale social media analysis during disasters and events more generally. The participatory sensing involves submitting flood reports and photographs and interacting with local authorities via a mobile app. The photographs are used to calculate river levels and flows using computer vision. The activity building involves engaging citizens based on their daily activities, i.e. cycling, walking, etc. and their location. This allows the authorities to request information in certain areas (which are managed by a control room to improve situation awareness) but also allows for narrowcasting of warnings, e.g. if driving, avoid a certain route due to flooding. The social media analysis allows authorities to spot emerging trends and pinpoint hotspots where activity is taking place. This application has been used most successfully during large social events such as festivals.

The second presentation related to environmental monitoring was on data fusion techniques for real-time mapping of air quality. Phillip Schneider showed how the vast quantities of point-based air quality measurements collected via their network (mostly set up in schools for educational outreach) could be used to update base maps of air quality generated from air dispersion models. This fusion allows for near-real time hourly updated information on air quality presented as a continuous field. These air quality maps can then be used to create products that are of value to citizens, e.g. calculations of the greenest route via bicycle from work to home or the level of personal exposure to certain pollutants in a given location. This information can also be of value to data from Sentinel-5P, i.e. for validation purposes or to downscale the data (which is at a 7km resolution) to finer urban scales.

1.4 Citizen Science to Support Conservation Management

Citizen science is most prevalent in the fields of nature conservation, ecology and biodiversity where the presentation by Eleanor Bruce clearly falls. Her area of interest is understanding the uncertainty in data collected by citizens in general but applied here to mapping species distributions of Humpback Whales. Since much of the citizen contributed data are opportunistically collected through tourism, the main concern in using the data for further scientific research is spatial bias. Using spatial analysis, her results showed locations where oversampling is most likely to occur and how this can be corrected, i.e. through better training of the citizen observers. Moreover, she reiterated the need to develop tools for communicating the uncertainty in the data collected by citizens.

1.5 Tools to Support Citizen Science

Zooniverse, as presented by Brooke Simmons, is a highly successful software platform for citizen science. Started in 2007, it currently has 1.4 million volunteers. A tool is now available at zooniverse.org/lab in which anyone can build a citizen science project on the Zooniverse platform. The guiding design principles of Zooniverse are embedded into this build-it-yourself product. Since starting, there have been 1,000 projects created, where around 5 will become part of the Zooniverse community and therefore have access to the Zooniverse volunteers. Zooniverse are happy to incorporate EO tools into the platform once they better understand what the needs of the EO community are.

Some DIY tools were demonstrated by Thomas Bartoschek of Münster University. There are two versions of the tool. The first is senseBox: home, an inexpensive DIY kit for building a sensor station to measure temperature, humidity, air pressure, loudness, illuminance and UV light with no technical expertise required. The sensor station can then be connected to the internet and the data are published on the sensemap website. The second version is for educational use (senseBox: edu), where the overarching aim is to teach students to build sensors, to learn how to program, to collect the data with their own sensors and then answer their own research questions. Future developments include sensors for air and water quality. DIY provides motivation and empowers citizens to ask questions about their environment, which they can then measure using these tools. As the data and technology are completely open, this provides easy access to development of community-based open science projects.

1.6 Issues Raised during the RoundTable

The following issues were raised and discussed during the roundtable:

1. How can we share data contributed by citizens in a standard format such as a WMS or WFS so that others can put all the pieces together and do further analyses? Stuart Wrigley answered that all EU-funded citizen observatories make the data available through GEOSS while Phillip Schneider said that all citizen-contributed data in Citi-sense were available by WFS. Alejandro Sánchez de Miguel raised the issue of costs involved in making the data available and suggested that GEOSS should also provide storage space. Heather Leson felt that there were broader issues beyond sharing of data, i.e. discoverability of research and how best to ensure this. Thomas Bartoschek provides all the data from their project via API, which he argued is more general than a WMS or WFS. This also ensures that wider audiences, i.e. journalists, can access and analyze the data. A member of the audience recommended that all the projects presented in this Citizen Science session should make it much clearer, on their respective websites, how the data can be obtained and the licensing.
2. Will automated methods replace the need for citizens in the future? Brooke Simmons of Zooniverse said that they are trying to build training sets using human computing in order to train automatic classifiers but that we will never be free from the need for human involvement. No matter how good training datasets are, there will always be new or rare objects that will not be picked up by an automated approach. Steffen Fritz added that there will be 9 billion human sensors in the world by 2050 and this human computing is irreplaceable.
3. Under what conditions does citizen science make sense, e.g. can the geolocation of ISS photos not be done automatically? Alejandro Sánchez de Miguel provided some technical reasons for why this

is a difficult problem and requires citizens, but he also made the point that with a press release and one month of time, all the photographs were geolocated. Hence the effort involved in motivating citizens to complete this task was not that large.

4. Do students (or people with more scientific background) perform better than citizens more generally in relation to Geo-Wiki and Picture Pile? Steffen Fritz answered that it is difficult to compare since the tasks given to the students in the past were much more difficult than the classification tasks in Picture Pile. However, these types of research questions could be answered in the future once the data are collected.
5. There is a massive potential to use citizens in increasing calibration and validation data needed for Sentinel, e.g. in identification of land cover, as LUCAS data are sub-optimal but represent one of the only sources for CORINE land cover. Are there are other examples that the presenters could think of? Frans Sniks said that we should think beyond just calibration and validation and look at the potential of higher spatial and temporal resolutions of measurements collected by citizens, e.g. information on aerosols. Alejandro Sánchez de Miguel made the observation that there are multiple apps that appear to do the same thing and that some apps could be adapted for other uses, e.g. the iSPEX app could be used for studies of energy efficiency. In other words, we should be joining efforts more closely. Brooke Simmons asked whether calibration and validation is standardized and requires the same training sets – if so, then she thought a simple project could be devised around this need.

2 Recommendations for Use of Citizen Science in EO Open Science

Based on the presentations, the roundtable discussions and citizen science more generally, the following recommendations can be made regarding future ESA projects and ESA's Open Science research agenda:

1. There is clearly a massive potential in involving citizens in the collection of calibration and validation data for EO, both in terms of image interpretation online and in-situ data collection via mobile apps. ESA should build on the inventory of existing citizen science initiatives, which have been mapped onto satellite sensors and products as started in the ESA-funded EducEO project, and then determine where gaps exist. Small pilot projects could be funded to fill these gaps.
2. Where existing projects and tools exist, ESA should work with these initiatives to test the use of these tools and data collected for calibration and validation of sensors/products. For example, the Picture Pile tool or Zooniverse platform could be used for collecting training data for a computer algorithm (e.g. computer vision) to classify satellite data automatically; the LACO-Wiki tool could be used for land cover validation; the community apps developed as part of the WeSenseIt project could possibly be used to generate training data for radar images such as Sentinel 1; the tools being developed in the Crowd4Sat project can be used for validation of Sentinel 1 and CORINE land cover; research could possibly be conducted to understand to what degree social media analysis can be used to support cal/val activities; different image classification projects could be added via the build-it-yourself project of Zooinverse; the Cities at Night project could be supported for applications relevant to ESA, etc.
3. Synergies between the projects presented in the Citizen Science session should be identified, e.g. the use of iSPEX for energy efficiency, the use of Picture Pile to help in humanitarian applications, etc. and then be supported as potential pilot projects in the future.

4. Resources should be invested in human computing to feed automated classification and computer vision algorithms in areas of relevance to ESA, e.g. classification of geotagged photographs that could become a source of calibration and validation data for land cover / land use maps.
5. More interactive tools and training materials are needed that are aimed at citizens and the next generation of users of EO, in particular for people working in low bandwidth environments. Visualization and communication of data uncertainty of citizen-contributed data should also be part of these tools as well as data quality more generally.
6. More events are needed in which the EO community interacts directly with the citizen science community so that that latter community can better understand the needs of EO providers and developers of EO products and services.
7. UAV imagery is increasingly being collected as part of post-disaster assessment. More demonstration projects are needed to realize the full potential of this data source for humanitarian purposes but also other calibration and validation activities.
8. ESA should make very high resolution Pléiades satellite imagery more accessible and/or use this data stream more effectively to engage citizens since this imagery is very attractive and easy to label. There are already applications that involve citizens in applications of food security, e.g. field size delineation, detection of water points, deforestation monitoring, etc. More demonstration project are needed to illustrate the untapped potential of crowdsourcing in this area.
9. To help mainstream citizen science into ESA activities, all future projects should have a component that considers how citizen science could be used to support the project or directly incorporates citizens in the data collection, analysis or educational outreach.
10. Demonstration projects are needed that link DIY measurements such as those collected through the sensor boxes demonstrated by Thomas Bartoschek with official measurements and EO data.
11. ESA should be a driver behind better data sharing of citizen-contributed data of relevance to EO or a major player in such an initiative. There are successful examples such as GBIF (Global Biodiversity Information Facility) and NatureServe (both for biodiversity data) that could be exemplars for such an approach. Although GEOSS represents a data brokering and discovery tool for shared EO data, it does not yet serve the needs of an EO-citizen science community.

3 Relevant References and Links

3.1 Relevant Publications

Bruce, E., Albright, L., Sheehan, S., Blewitt, M. 2014. Distribution patterns of migrating humpback whales (Megaptera novaeangliae) in Jervis Bay, Australia: A spatial analysis using geographical citizen science data. Applied Geography, 54, 83-95.

Clery, D., 2011. Galaxy Zoo volunteers share pain and glory of research. *Science*, 333 (6039), 173–175.

Fonte, C., Bastin, L., See, L., Foody, G. and Lupia, F. 2015. Usage of VGI for validation of land cover maps. *International Journal of Geographic Information Science*. http://dx.doi.org/10.1080/13658816.2015.1018266

Kyba, C., Garz, S., Kuechly, H., de Miguel, A., Zamorano, J., Fischer, J., and Hölker, F., 2014. High-resolution imagery of Earth at night: New sources, opportunities and challenges. *Remote Sensing*, 7 (1), 1–23.

See, L., Fritz, S., Dias, E., Hendriks, E., Mijling, B., Snik, F., Stammes, P., Vescovi, F., Zeug, G., Mathieu, P.-P., Desnos, Y.-L. and Rast, M. A new generation of tools for crowdsourcing and citizen science to

support Earth Observation calibration and validation. In revisions. *IEEE Geoscience and Remote Sensing Magazine*.
See, L., Fritz, S., Perger, C., Schill, C., McCallum, I., Schepaschenko, D., Karner, M., Kraxner F., Obersteiner, M. 2015. Harnessing the power of volunteers, the Internet and Google Earth to collect and validate global spatial information using Geo-Wiki. *Technological Forecasting and Social Change*. doi:10.1016/j.techfore.2015.03.002
See, L., Perger, C., Hofer, M., Weichselbaum, J., Dresel, C. and Fritz, S. 2015. LACO-Wiki: An open access online portal for land cover validation. ISSDQ, Montpellier, 29-30 Sep 2015.
Snik, J. H. H. Rietjens, A. Apituley, H. Volten, B. Mijling, A. Di Noia, S. Heikamp, R. C. Heinsbroek, O. P. Hasekamp, J. M. Smit, J. Vonk, D. M. Stam, G. van Harten, J. de Boer, C. U. Keller, and 3187 iSPEX citizen scientists. 2014. Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters, *Geophys. Res. Lett.*, 41(20), 2014GL061462.
Sauermann, H. and Franzoni, C., 2015. Crowd science user contribution patterns and their implications. *Proceedings of the National Academy of Sciences*, 112 (3), 679–684.

3.2 Relevant Platforms and Websites

Relevant platforms and websites that were mentioned during the Citizen Science session include:

- Crowdsourcing of land cover via Geo-Wiki: http://www.geo-wiki.org
- Picture Pile serious game for land cover classification: http://www.geo-wiki.org/games/picturepile/
- Validation of land cover: http://www.laco-wiki.net
- MicroMappers: http://micromappers.org/
- Humanitarian OpenStreetMap: http://hotosm.org/
- EU citizen observatories: http://www.citizen-obs.eu/
  - WeSenseIt: http://wesenseit.eu/
  - Citi-Sense: http://www.citi-sense.eu/
- Zooniverse: https://www.zooniverse.org/ -build your own project: https://www.zooniverse.org/lab/
- iSPEX: http://ispex.nl/en/
- SenseBox: http://sensebox.uni-muenster.de/en and OpenSenseMap: http://opensensemap.org/
- Cities at Night: http://www.citiesatnight.org/ and processed data at : http://pmission.cartodb.com
- Crowd4Sat: http://www.crowd4sat.eu/

Other relevant citizen science platforms and websites that could be used as sources of calibration and validation for EO (extracted as part of the review undertaken in the ESA-funded EducEO project):

- Public Laboratory Balloon and Kite Mapping: http://publiclaboratory.org/tool/balloon-mapping
- Omniscientis: http://www.omniscientis.eu/
- The NOVA Labs: Cloud Lab and Sun Lab: http://www.pbs.org/wgbh/nova/labs
- Students' Cloud Observations Online (S'COOL): http://scool.larc.nasa.gov
- SatCam: http://satcam.ssec.wisc.edu
- Cobweb (Citizen Observatory Web): http://cobwebproject.eu
- Global Learning and Observations to Benefit the Environment (GLOBE): http://globe.gov
- FieldScope: http://education.nationalgeographic.com/education/programs/fieldscope
- The Almanac: http://www.thealmanac.org/
- SkyWarn: http://www.skywarn.org/)
- Did You Feel It?: http://earthquake.usgs.gov/earthquakes/dyfi/

- Did You See It?: http://landslides.usgs.gov/dysi/form.php
- Mount Diablo Fire Monitoring: http://nerdsfornature.org/monitor-change/diablo.html
- Frack Finder: http://crowd.skytruth.org/
- Field Photo Library: http://www.eomf.ou.edu/photos/
- Degree Confluence Project: http://www.confluence.org/
- International Network of Crisis Mappers: http://crisismappers.net/
- Ushahidi: http://www.ushahidi.com/
- TOMNOD: http://www.tomnod.com/nod/challenge/
- The Digital Humanitarian Network: http://digitalhumanitarians.com/
- Dark Sky Meter: http://www.darkskymeter.com/
- Aurorasaurus: http://aurorasaurus.org/
- Citclops: http://www.citclops.eu/
- USA National Phenology Network: https://www.usanpn.org/
- Public Laboratory Infrared Camera: http://publiclaboratory.org/tool/near-infrared-camera
- CoCoRaHS: Rain, Hail, Snow Network: http://cocorahs.org/
- Rainlog.org: http://rainlog.org
- Tracking Climate in Your Backyard: http://www.priweb.org/outreach.php?page=citizenscienceed/TCYIB
- Precipitation ID Near the Ground (PING): http://www.nssl.noaa.gov/projects/ping/
- Snow Tweets: http://www.snowtweets.org/
- Our Radioactive Ocean: http://ourradioactiveocean.org/helpus.html
- Deforestation Mapping in Canada: https://cfsnet.nfis.org/deforestation/
- Forest Watchers: http://forestwatchers.net/
- Treezilla: http://treezilla.org/
- Urban Forest Map: http://urbanforestmap.org
- Urban Tree Survey: http://www.nhm.ac.uk/urban-tree-survey
- EarthWatchers: http://dfa.tigweb.org/)
- Weather Underground: http://www.wunderground.com/
- WOW: http://wow.metoffice.gov.uk/
- Citizen Weather Observer Program: http://www.wxqa.com/