CROWD4SAT

_____

Using crowd-sourcing and citizen science techniques with Observations from Satellites

Crowd4Sat-DEL-2.2

FINAL REPORT

Deliverable 2.2 / Deliverable 3.3

*Utility Assessment Report
Final Report*

May 2017

Issue 1
Revision 3e

Imperative Space

Contact: Ravi Kapur
             Managing Director
             Imperative Space

        tel: +44 (0)1234 867115
       mob: +44 (0)7977 412739

     e-mail: ravi.kapur@imperativespace.com

1 Executive Summary

1.1 Introduction

The Crowd4Sat initiative led by Imperative Space has spanned a wide range of subject matter in Earth observation and astronomy, to test the opportunities for using citizen science and crowdsourcing (CS) in relation to data originated from satellite platforms, and to develop new workflows and techniques for rapidly incorporating data into CS platforms and generating useable results. Our particular CS approach focussed predominantly on the use of 'human pattern recognition', utilising established image sorting and classification platforms. These platforms have in turn benefited from considerable extension to their functionality as a direct result of the research activity and software development conducted through Crowd4Sat.

In particular, the Zooniverse platform now has the capability to handle Earth observation data in a way that did not previously exist, and these new features have also fed into other classification projects on the platform. The Picture Pile platform, (used in the humanitarian-focused contract extension to the project), has also benefited from such development work. Full details of the new features and specifications enabled through this project are provided later in this report.

This report synthesises previous reporting from the various strands of the project and incorporates more recent analysis and overarching commentary. It also references initial outcomes from the humanitarian contract extension project. Full results from that strand will be included in a separate report specific to that project.

In broad terms, Crowd4Sat can be seen as having produced successful outcomes in terms of the research and development activity to assess, test and analyse the specific ways in which citizen science and crowdsourcing (through human pattern recognition activity and related software elements), can be harnessed and deployed. It has shown how these tools can be used to great effect in helping to reach scientific, conservation and research goals, and as a part of data validation.

However, in the case of the Earth observation strands of the project, the extent of the 'real-world' outcomes and involvement of significant numbers of citizen scientists has been somewhat limited due to a number of inherent technical challenges, (which in fact proved to be a key focus of the research effort in the project). Solutions to these challenges have been identified during the course of this project, but more detailed development will still be required in future projects before meaningful real-world deployments of this kind can be made routinely possible.

This largely relates to issues of resolution of the satellite data, workflows for image processing and correction, and the need to develop new methods for providing training data and guidance to citizen scientists. These points, and a description of potential follow-up projects which may be required to address them, are dealt with later in this report.

This report consists of the following elements and appendices:

- This executive summary.
- Summary of new feature development on the Zooniverse platform, enabled through the Crowd4Sat project.
- Summary of outcomes from core demonstration projects.

- Summary of 'challenge project' experimental pilots, their challenges and research outcomes.
- Outline framework for a proposed rolling 'Environmental Tagging Service'
- Summary of recommendations and outline 'roadmap'.
- Notes on potential citizen science-based EO services.
- Appendix 1 – Additional screenshot images from the pilot projects.
- Appendix 2 – Foresight report (D1.1) (with brief addendum).
- Appendix 3 – Strategic roadmap (D1.2) (cross referenced from new notes in this report).
- Appendix 4 – Pilots requirements specification (D1.3 + D2.1) (cross referenced from new notes in this report).
- Appendix 5 – Interim report on humanitarian CCN (final report expected by late-July/early August 2017).
- Appendix 6 – Summary consultation report on opportunities with humanitarian NGOs (conducted late 2015).

1.2   Summary of research activities and main outcomes

1. The Crowd4Sat initiative set out to test the effectiveness and usefulness of citizen science and crowdsourcing methodologies in the processing, analysis and validation of data obtained from satellite platforms. This encompassed astronomical observations from space-borne telescopes as well as Earth observation data. The consortium led by Imperative Space focussed predominantly on a particular type of citizen science activity, involving the use of human pattern recognition capability to rapidly classify large numbers of images. This methodology is well established, particularly in the astronomy arena, and our main platform partner in the consortium ('Zooniverse', based at the University of Oxford) is a recognised world leader in this approach.

2. The subject areas for the demonstration projects focussed mainly on astronomy, conservation and the use of Earth observation in humanitarian activity. An additional strand of activity in the humanitarian area, using an alternative classification platform called 'Picture Pile' (developed by IIASA), was also enabled through a contract extension project. Three so-called 'core' projects utilised the Zooniverse citizen science image classification platform developed by the Department of Astronomy at the University of Oxford. These projects were designed to be delivered at scale, tapping into the existing global Zooniverse user-base of committed and enthusiastic citizen scientists around the world. Additional 'challenge' projects also utilised the Zooniverse platform, but in a different method of deployment designed for smaller or experimental projects.

3. The Zooniverse team has extensive experience of engaging audiences of tens or hundreds of thousands of people on individual classification projects, and has a deep understanding of the motivations with which citizen scientists engage in sometimes mundane and repetitive micro-tasking, in the belief that they are contributing to important scientific research or socially valuable outcomes.

4. The astronomy projects originally planned for this initiative were focused on observations of star formation using the Herschel mission, and supernovae using the Gaia mission. In practice, delays and changes in the scientific research teams meant that some changes to these demonstration projects had to be made. Whilst a preliminary test project on star formation was conducted using ground-based data, the Herschel pilot was in fact replaced

by an alternative pilot project looking at measurements of galaxy 'bar lengths' using Hubble data. Further details of this are provided later in this report. (It was accepted that the Herschel project was no longer deliverable due to factors within the Herschel research team outside of our control. The replacement project used Hubble data, and was in effect the second phase of a project which was built and run previously (outside of Crowd4Sat funding) but now incorporating some features and code developed through Crowd4Sat).

5. Both of these astronomy projects resulted in significant findings and involved many thousands of users and tens of thousands of classifications in a relatively short period. Some scientific publications have also resulted from this work. Full outcomes are provided later in this report.

6. The third of the 'core' pilot strands focussed on a novel approach to utilising Earth observation data to assist first responders when activated in response to humanitarian crises. This involved use of Sentinel-2 data, and a wider collaboration with other data providers (including Planet Labs), a first responder NGO and other expert advisers in the humanitarian field. In the early research and development phase of the project, an opportunity arose to run an initial test pilot project, in response to the Nepal earthquake in 2015. This pilot was conducted somewhat 'under the radar' with a closed group of classifiers from partner organisations and first responder NGOs. The outcome of this spontaneous pilot project showed some anecdotal evidence of imagery classified through the CS platform being used in on-the-ground response activity in Nepal. In addition, the process of finding and processing the relevant imagery provided lessons which fed into the more substantive disaster response project conducted later in the initiative.

7. The later project in the humanitarian 'core' strand was conducted in a real-world context in response to the Ecuador earthquake in April 2016. The citizen science platform, workflow testing, and collaborative arrangements with providers had been established just in time to enable the deployment. This deployment also put in place the so-called 'Planetary Response Network' (PRN) website and partnership. The PRN concept is outlined later in this report. (However, this partnership has not yet been taken further forward as core funding is still required to enable the partnership to be maintained).

8. Overall, the three core projects using the Zooniverse platform involved tens of thousands of classifiers (citizen scientists) and led to tangible, published scientific outcomes, and numerous software developments and new workflows. It has also led to further exploration and discussion about the value of such methodologies in relation to humanitarian response amongst relevant organisations.

9. In parallel to the three 'core' demonstration projects, the consortium undertook additional research to explore the opportunities for novel uses of citizen science in slightly more experimental contexts, with a particular focus on using Sentinel-2a data. This again involved use of the Zooniverse platform and a human pattern recognition image classification approach, but with the added benefit of new tools and features added to the platform as a direct result of the Crowd4Sat funding. A range of potential demonstration projects (so called 'challenge' projects) were explored and co-developed with specific end-user partners, including WWF and advisors from King's College London. (A list of these initial ideas can be found in the earlier project specification documents in Appendix 4).

10. After conducting research on the viability of running challenge pilots within the timeframe available and within the partnerships established for the project, the subject areas eventually settled on were as follow:

- Monitoring of waterholes to support conservation activity for WWF
- Monitoring illegal gold mining activity in French Guiana, again with WWF
- A project in partnership with Earth observation and land cover specialists at Kings College London to explore the possibility of creating a citizen science-based data validation tool for existing forest fire-related data products.

11. Again, the process of researching, setting up and creating new infrastructure and platform features to enable the challenge pilot projects has provided considerable learning and tangible software outcomes. In the case of these challenge projects, it is this learning and the workflows established which have been the key beneficial outcomes to have emerged, rather than research outcomes from the classification process itself. Further summaries, screenshots, and context for the challenge projects are provided later in this report.

12. However, the challenge projects did not prove to be as successful as hoped as actual real-world deployments of citizen science (in the way that the core projects did). The principal reason for this was the mismatch between the resolution of the data available from Sentinel-2a and the level of detail required to meet the aims of the challenge projects. A more in-depth explanation of this is provided in section 4 below.

13. In addition to development of the demonstration projects, work was also undertaken to develop the ideas and framework for a potential future CS-based 'environmental tagging service' to enable easier management of appropriate data required for environmental research, campaigns, conservation, and NGO activity. A summary of the ideas and potential features for this service are outlined in section 6 below, and are also linked to the 'Recommendations and Roadmap' also outlined later in this report. In practice, as indicated in the recommendations, a dedicated additional project will be required to take this concept to the next stage.

14. Also included as an Appendix to this report is the interim report from the separate humanitarian contract extension strand, which is still currently in progress. Further pilot work and analysis is due to be completed in June and July 2017, and a full report for this strand will be issued separately in due course.

1.3   Summary of demonstration projects conducted

The final list of demonstration pilots under this project evolved from the original plan in response to a number of factors, including availability of adequate data, guidance from specialist partners, and technical limitations. These changes were discussed with ESA at key stages.

The pilots were grouped under four themes: 1) Astronomy; 2) Humanitarian; 3) Wildlife and Biodiversity; 4) Environmental Monitoring. However, practical delivery of the projects was divided into two groups: 'Core' projects and 'Challenge' projects. These are summarised as follow:

'Core' projects:

1) Planetary Response Network – Sentinel 2 and Planet Labs data (humanitarian)
2) Galaxy Zoo: Bar Lengths – Hubble data (astronomy)
3) Project Gaia (classification of transients) – Gaia data (astronomy)

'Challenge' projects (beta tests):

1) Wildlife Waterholes – in partnership with WWF, using Sentinel 2 & Urthecast data (wildlife & biodiversity)
2) Guiana Goldmining Watch – in partnership with WWF, using Sentinel 2 & Urthecast data (environmental monitoring)
3) Burn-Scar Mapping validation – in partnership with advisors from KCL, using Sentinel-2 data (environmental monitoring)

Additional work-streams:

- Development of a framework and ideas for an on-going 'environmental tagging service'.

- Development of a comprehensive set of new features for the use of EO data on the Zooniverse platform (guided by and specified in consultation with Imperative Space).

- Research and consultation with NGOs on the potential for citizen science-led projects to support humanitarian response and planning activity.

- (The earlier research phase of the project also included a detailed literature review, analysis of CS projects and platforms, and exploration of the educational value of CS).

Humanitarian CCN:

- A separate strand of activity was instigated under a contact extension, to explore specific approaches to image classification in response to humanitarian crises, using the IIASA/GeoWiki 'Picture Pile' platform, and in partnership with the Humanitarian Open StreemMap Team and MapSwipe. Further information about this is provided later in this report and in Appendix 4.