Vol-xxxx
urn:nbn:de:0074-xxx

ENVIP'2015 (DRAFT)

Environmental Infrastructures and Platforms 2015 - Infrastructures and Platforms for Environmental Crowd Sensing and Big Data

**Proceedings of the Workshop " Environmental Infrastructures and Platforms 2015 - Infrastructures and Platforms for Environmental Crowd Sensing and Big Data
co-located with the European Citizen Science Association General Assembly 2015 (ECSA GA'2015)**

Barcelona, Spain, October 28-30 October, 2015.

Edited by

Arne J. Berre *
Sven Schade **

Jaume Piera ***

* SINTEF, Oslo, Norway
** JRC , Joint Research Centre, Ispra, Italy

*** ICM-CSIC , Institute of Marine Sciences, Barcelona, Spain

Table of Contents

Harnessing location-based services for effective Citizen Observatories

Suvodeep Mazumdar[1], Stuart N. Wrigley[1], N.Ireson[1], F.Ciravegna[1]

[1]Organisations, Information and Knowledge (OAK) Group,
Department of Computer Science,
University of Sheffield,
211 Portobello, Sheffield, UK
{s.mazumdar,s.wrigley,n.ireson,f.ciravegna}@sheffield.ac.uk

Abstract

The essence of a city is its citizens and communities. A city's infrastructure and associated services play a vital role in citizens' day to day living and their overall quality of life. Traditionally, services are deployed in a top-down approach where authorities, councils and public bodies take a reactive approach to address community needs and concerns. In this paper, we propose our 'Citizen Observatory' approach to enable citizens to take a pro-active role in the management of their local communities and environment by supporting their engagement in the decision-making process. We discuss how to empower citizens and communities to engage with and assist authorities to establish a more informed understanding of residents' needs and the status of their local environments. Through the WeSenseIt project, we employ a location-based crowdsourcing and communication strategy to develop a resilient, efficient and collaborative information ecosystem for decision-making in city, urban and rural areas.

Keywords: Citizen Observatories, Geofencing, Crowdsourcing

1. INTRODUCTION

Citizens and communities are traditionally considered as mere `passive' consumers of services at the very end of the information chain. In an ever-changing technological landscape this view is, however, continuously challenged. Citizens and communities demand and expect newer forms of communication and democratic processes to aid further engagement with policy makers. While there is a greater need for transparency, engagement, collaboration and information sharing, traditional approaches are increasingly unable to address such expectations. We aim to address this gap by employing Information and Communications Technologies (ICT) to enable new forms of interaction, knowledge exchange and participation of the general public, decision-makers and experts. This creates a shift from a traditional one-way communication paradigm (where organisations gather information from internal or trusted sources and communicate this to citizens) towards a two-way collaborative environment: citizens become active members with respect to the provision of information and authorities have a duty to consider, respond and react to citizens' communications. Public engagement can increase democracy (sharing of information); broaden community/public education; provide local data, knowledge and problem-solving; monitoring situations that otherwise would not be; and potentially provide early warning/detection systems for emergencies (Conrad et al, 2011).

Informal social networks and Web 2.0 technologies have already facilitated wide scale adoption of crowdsourcing in which citizens act as information providers across a broad range of domains (from collaborative traffic and navigation to environment monitoring (Heipke et al, 2010)). The proliferation of Web 2.0 services and applications, along with 'always-connected' mobile technology has paved the way for highly successful crowdsourcing-based knowledge creation ventures. Wikipedia, Flickr and OpenStreetMap serve as excellent real-world examples where crowdsourcing has provided immense wealth of information to be used by organisations and communities world-wide. In fact, most popular websites such as TripAdvisor, Amazon, eBay and other e-commerce platforms exploit the potential of crowdsourced data to provide customers with a greater understanding of the value of the purchases they intend to make. While such ventures have been proven to be highly successful in websites and applications, employing crowdsourcing for city-scale planning and policy-making is a more complex process. The real potential of crowdsourcing for governance, policy-making and urban planning can only be realised when opinions, views and concerns are shared between citizens, authorities and decision-makers in a collaborative and seamless manner. True participation therefore requires active engagement of the public in planning and decision-making and as such relies on communication and knowledge exchange.

In the WeSenseIt project, we employ various forms of crowdsourcing techniques designed to empower and foster participation with the objective of creating a highly enriched, real-time knowledge base to aid decision-making processes focussed around water and flood management. Our approach facilitates citizen cooperation, engagement and participation in various forms of decision-making. We term this framework, from a holistic point of view as a `Citizen Observatory' (CO). While several definitions exist, we formally define a citizen observatory as ``a method, an environment and an infrastructure supporting an information ecosystem for communities and citizens, as well as emergency operators and policymakers, for discussion, monitoring and intervention on situations, places and events" (Ciravegna et al, 2013). COs are enormous resources for data collection and classification at a granularity and frequency that is unreachable with traditional methods. The majority of CO and citizen science projects stem from scientists and authorities that have strong organisational support. Research has shown that crowdsourced data can equal or even exceed that produced by professionals (Zaidan et al, 2011). COs are emerging as a means to establish interaction and co-participation between citizens and authorities during day-to-day management of fundamental resources (Lanfranchi et al, 2014).

This paper presents how the WeSesenseIt project employs crowdsourcing and traditional mechanisms to gather information from sensors and citizens. Additionally, we present our location-based geofencing approach to solicit and share information. In our implementation, a geofence is a virtual enclosed area (bounded by a regular or irregular polygon), entering which triggers an alert on a citizen's mobile device to inform of potential hazards or seek local information, media or data. The next section presents our two-layered approach for sensing data in an urban environment. Section 3 then presents our geofencing approach for soliciting information from citizens. Section 4 briefly discusses how the information gathered is analysed within the WeSenseIt framework and Section 5 concludes the paper by discussing evaluation plans and future work.

2. SENSING THE PULSE OF THE CITY: THE WESENSEIT APPROACH

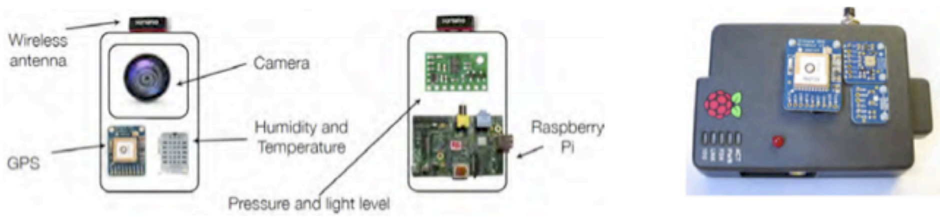
Our Citizen Observatory model incorporates two layers of sensing to enable authorities gain a multi-faceted insight into the city: hard and soft layer. The hard layer includes static and portable devices and sensors that sense physical quantifiable parameters such as water levels, soil moisture, occupancy etc. These devices are either installed as professional sensing equipment in weather stations or observation points, or built-in within portable devices controlled by or resident within citizen's mobiles or tablets (e.g. pervasive sensing). The soft layer aims at harnessing the `collective intelligence' of citizens: this is gathered from purpose-built mobile and web-based applications; harvested from social networking platforms (e.g. social sensing); and the wider web. The ability to `listen' to citizens is the first step in developing a collaborative framework for a citizen-centric smart city. The soft layer also provides `hard data', where citizens can provide raw readings as observed from a sensor (e.g. the room temperature, as read from an analogue unit). The hard and soft layers complement (and correlate with) each other, providing two different views on an observed phenomenon. As a result, this provides a greater understanding of the spatial and temporal evolution of human behaviour based on various observed phenomena.

2.1. The Hard Layer

Two main types of sensors dominate the hard layer: professional sensors and pervasive sensors. In our approach, the hard layer provides data about phenomena that are measurable and quantifiable. Examples of such devices are weather stations measuring wind speed, wind direction, precipitation etc. The professional sensors are high quality and high precision sensors installed at very few locations across cities. While in an ideal scenario such sensors should highly distributed within the city area, the high costs are often prohibitive. Such sensors are therefore installed sparsely across large geographical spaces, in high priority locations as defined by authorities. Pervasive sensors, on the other hand, are more economically viable but at the cost of reduced quality and precision. Despite of the lower precision and data quality, a large number of pervasive sensors can be deployed by authorities and citizens that can provide a highly granular set of data. The high coverage of the sensors, viewed in unison can provide a significant insight into the phenomena in consideration. The hard layer therefore consists of:

- A set of traditional, professional-grade sensors and weather stations designed and installed for remote sensing of physical variables. These are installed in high priority locations, providing a stream of high quality data.;
- A set of innovative low-cost sensors (Figure \ref{fig:hardlayer}) for reading a variety of phenomena and physical variables such as water levels. These sensors are available as pre-installed, pre-programmed units or as design plans. Enthusiasts can use the design plans to build their own sensors, with software easily available. Citizens can then install such sensors in their own premises and start providing data to the platform.;

Figure 1: Examples of low-cost sensors provided to citizens for installation. Image from Lanfranchi et al 2014



2.2. The Soft Layer

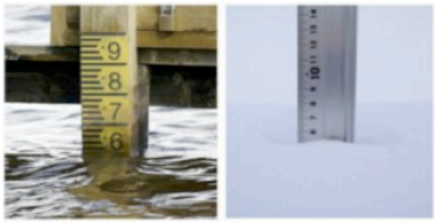
The soft layer aims to generate data directly from citizens' activity. Based on the amount of activity, there are two types of soft layer sensing: active and passive. Active sensing requires citizens to actively provide direct feedback on topics of interest: authorities can contact citizens and request them to provide specific information. One of the most well-known examples is the FixMyStreet app, where citizens can submit a geo-located form to report issues such as street cleaning, lighting issues, etc. On the other hand, passive sensing provides a means to understand the user's opinions and perceptions based on their activity elsewhere. For example, a user commenting on social media that they are upset due to a delayed bus service provides feedback on the operation of a bus service, while the original intent was to express disappointment and frustration for the user being delayed for their next trip. The soft layer consists of:

- A set of mobile applications that support communication and information sharing. Citizens can receive critical information and alerts to be informed of situations around them. They can also provide physical readings of sensors or observations. The applications also provide citizens with means to comment, report or highlight their concerns.
- A set of Social Media analytics technologies. These tools reside in the background and can be used in multiple ways: proactively seeking information related to critical issues reported by citizens or trigger alerts when urgent scenarios emerge (e.g. reports of earthquakes or house fires in neighbourhoods)

3. GEOFENCE-DRIVEN CROWDSOURCING

Professional and pervasive sensing technologies provide a large amount of information on areas covered by sensors. The sensors are largely limited to static locations pre-determined by authorties or citizens based on an information need or citizen-interest. Despite the availability of cheaper sensors, obtaining higher levels of coverage in a large geographical area is still a highly complex and expensive process. This is further complicated with the need for such sensors to be maintained (i.e. batteries, occasional restarts, vandalism, herds/cattle disturbing alignments etc.) particularly if installed in open spaces as opposed to private properties under the attention of a citizen. Furthermore, in a city-wide infrastructure, authorities are challenged with highly evolving and dynamic scenarios every day. For example, large-scale events organised in cities see a massive movement of citizens across various areas. Daily commuting activities also exhibit wide variability in mobility where citizens travel to/from work at different/fixed times every day. Emergencies such as floods and fires require real-time understanding of evolving situations: different regions are affected due to various factors. For example, heavy rain in neighbouring areas, tidal forces, sustained periods of rains, flood plains overflowing etc. have different impacts on locations. In such cases, the area of interest changes dynamically and rapidly. Often, new areas of interest are not well-covered by existing sensor networks and hence authorities remain poorly informed regarding a rapidly evolving event in a ‘dark-spot’. Furthermore, it is common for authorities to install low-cost analogue sensing devices in a large number of locations (Figure 2); for example, rain/snow gauge boards, wind speed devices, etc. The issue with such units is that they are `dumb': they cannot transmit data autonomously and therefore require manual intervention to communicate their readings to help authorities understand local conditions. In a period of budget cuts and, in turn, falling staffing levels in authorities, it is increasingly common for authorities not to have the resources (e.g., staff numbers) to send employees to read and record the values of such analogue sensing devices.

Figure 2: Examples of analogue guage boards



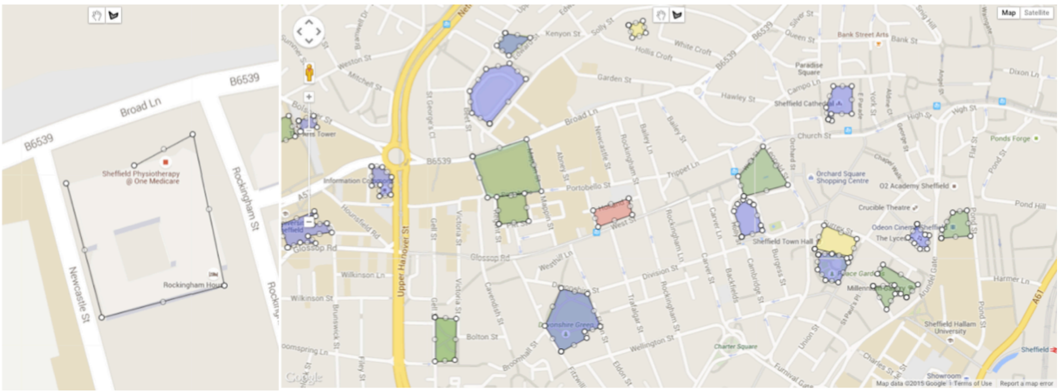
In urban environments, a large majority of locations are covered by citizens and communities. We aim to exploit this by empowering citizens with crowdsourcing techniques to expand to new regions as and when a demand for information is established. This provides authorities with a better understanding of new regions, supported by a collaborative partnership with citizens. Citizens can contribute significantly by acting as the `eyes and ears on the ground': directly informing authorities of the analogue sensor values, sending an image or video, streaming audio/video to control units etc. We describe our approach~\cite{mazumdar2014} from two perspectives:

3.1. Geofence Generation

Upon identifying the need for information related to a specific region in the city, the authorities create a geofence by drawing a polygon on a map. This prompts authorities to define the kind of geofence (area of interest, danger area, safe area, historical etc.) the area represents. Authorities can further declare tasks and actions associated with each geofence, with a period of validity. Upon completion of the geofence definition, authorities can view the different types of geofences they have created and can finally publish them. Figure 3 shows a resulting screenshot of a geofenced region.

The image on the left of Figure 3 shows a user defining an area of interest (in this case, a building) by drawing a polygon on a map. Upon completing the polygon, the user provides further details of the geofence (name, description, type, associated actions). The various geofences are then colour-coded to quickly help interpret the various types of geofences. Red geofences represent unsafe zones and are marked as danger areas: in an emergency, such areas should be avoided by citizens. Blue areas represent areas where analog sensors are installed: citizens can provide data from these areas. Yellow areas indicate areas that are of historical significance. These categories of regions are often evolving and varied across different cities, use cases and needs.

Figure 3: Creating Geofences (left) using mouse gestures to define boundaries. Colour-coded geofences for different related tasks (right): red indicating danger, green indicating points of interest, blue indicating areas with analog sensors.



3.2. Citizen Response

Citizens are provided with mobile device applications that run in the background. These applications are a part of our typical `soft layer' of data collection which citizens can use to communicate with authorities. The `always-on' facility of the applications make timely queries to the geofencing services to understand if the citizen is present in a particular location. Upon entering a geofenced area, the device triggers an alert / notification. The citizen can choose to ignore or respond to the alerts accordingly. A variety of alerts are made available to citizens, based on their needs and the needs of the situations. For example, upon entering a danger area, a citizen is provided with a different alert as opposed to a historically significant area. In the case of the former, an alert like a notification message, vibration and sound would be more useful, while a simple popup to indicate an interesting `story' would be sufficient for the latter. Notifications and alerts can be managed by users, depending on their level of interest, and can also be `learned' based on their behaviour. If the citizen chooses to respond to the alerts, a variety of possible combinations are proposed based on the needs of authorities. Citizens can merely respond with a video/image of their surroundings such as taking a picture of the river. Alternatively, citizens could respond with an audio recording of what they can hear in their present location (if authorities are interested in understanding the level of noise generated in particular areas). Citizens can also provide analogue sensor readings if needed, by observing analogue sensing devices, typing the value into the app and submitting their report. Once the citizen responds to the notification, authorities are immediately updated.

4. REAL-TIME ANALYSIS OF CROWDSOURCED DATA

While gathering large volumes of data is a significant challenge, it is only one aspect of a citizen-authority information processing workflow. Massive streams of data arriving from a variety of users, data providers, sensors need to be processed in real-time to provide authorities with updated information. This is a critical need, since some scenarios may need urgent attention of authorities (e.g. emergencies). We approach this by conducting multi-level analyses on different streams of data. Real-time sensor data are processed to be stored immediately in large datastores, which are then subsequently used by analytics modules (e.g. visualisations and real-time monitoring systems). Citizen-generated data, on the other hand is processed based on their type. Named entities (locations, organisations, identities, names etc.) are extracted automatically from text entries such as comments, social media messages and form data. Sensor readings submitted by citizens are validated and stored in sensor data stores. Exif data from images is extracted to provide further metadata. Images are also tagged with any additional information that a citizen provides to compliment their observation. All information is stored in datastores, indexed and available to be quickly retrieved when needed.

5. CONCLUSIONS AND FUTURE WORK

In this paper we presented a two-layered approach adopted in the WeSenseIt project to gather data from physical and social sensors. We also presented our location-based geofencing approach that is aimed at soliciting and communicating critical information based on a user's location. Finally, we presented how the collected crowdsourced information is analysed to help decision makers and authorities take critical decisions. The WeSenseIt project is in its final stages and at the time of writing this paper, has just entered into its final evaluation phase. These technologies will be evaluated over the next few months in several exercises across Europe. A game-based evaluation exercise is also scheduled to be organised in Sheffield that will evaluate the geofencing approach, where volunteers will be provided with mobile applications and a goal of the game. Multiple areas of interest will be defined and users will be prompted to provide information when they enter such areas. The evaluation will attempt to understand how effective the geofencing approach is in providing information to authorities, based on dynamic geofences. Future planned studies also include an evaluation of the geofencing approach with a non-geofence based approach in a comparative setting. Finally, the geofencing application is planned to be distributed to wider populations to help citizens contribute local information to freely available open datasets, based on missing local information

REFERENCES

Conrad, C. C., and Krista G. H. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental monitoring and assessment* 176.1-4 (2011): 273-291

Heipke, C..Crowdsourcing geospatial data. ISPRS *Journal of Photogrammetry and Remote Sensing* 65.6 (2010): 550-557

Ciravegna, F., Huwald, H., Lanfranchi, V. Wehn de Montalvo, U. Citizen observatories: The WeSenseIt Vision. *Inspire Conference*, Florence, 2013

Lanfranchi, V., Wrigley, S. N., Ireson, N., Wehn, U., \& Ciravegna, F. Citizens' Observatories for Situation Awareness in Flooding. *Proceedings of the 11th International ISCRAM Conference*, University Park, Pennsylvania, USA, 2014

Mazumdar, S. Wrigley, S.N., Ireson, N. and Ciravegna, F. Geo-fence driven crowd-sourcing for emergencies. *Proceedings of the 12th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, 24-27 May, Kristiansand, Norway, 2014

Zaidan, O. F., Callison-Burch, C.. Crowdsourcing translation: Professional quality from non-professionals. *In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1 (pp. 1220-1229)*. Association for Computational Linguistics, 2011

Citizen OBservatory WEB (COBWEB): A Generic Infrastructure Platform to Facilitate the Collection of Citizen Science data for Environmental Monitoring

Christopher I. Higgins[1], Jamie Williams[2], Didier G. Leibovici[3], Ingo Simonis[4], Mason J. Davis[5], Conor Muldoon[6], Michael O'Grady[6]

[1]EDINA, University of Edinburgh, chris.higgins@ed.ac.uk
[2]Environment Systems Ltd., [3] Nottingham Geospatial Institute, University of Nottingham, [4]Open Geospatial Consortium, [5]Welsh Government, [6]University College Dublin

Abstract

COBWEB has used the UNESCO World Network of Biosphere Reserves as a testbed for researching and developing a generic crowdsourcing infrastructure platform for environmental monitoring. A major challenge is dealing with what is necessarily a complex problem requiring sophisticated solutions balanced with the need to present sometimes unsophisticated users with comprehensible and useable software. The components of the COBWEB platform are at different Technology Readiness Levels. This short paper outlines the overall solution and points to quality assurance, standardisation and semantic interoperability as key areas requiring further attention.

Keywords: citizen science, crowdsourcing, Open Geospatial Consortium, environmental governance, spatial data infrastructure, sensors, access control, privacy.

INTRODUCTION

New and innovative environmental monitoring and information capabilities can enable effective participation by citizens in environmental monitoring, based on broad stakeholder and user involvement in support of both community and policy priorities (Liu et. al. 2014). The objective of Citizen OBservatory WEB (COBWEB) has been to research and develop an innovative generic infrastructure platform to facilitate the collection of citizen science data for the purpose of such environmental monitoring and governance. With a particular focus on the use of open interoperability standards, COBWEB demonstrates how advances in mobile and sensor technology combined with the large increases in availability of mobile devices, especially of smartphones, can equip citizens to make observations of use for good environmental governance.

COBWEB has focused on three pilot case study areas: the creation and validation of data products from Earth Observation data; biological monitoring; flooding. To evaluate these case study areas, testbed environments have been established within the United Nations Educational, Scientific and Cultural Organization's (UNESCO) World Network of Biosphere Reserves. Modern Biosphere Reserves can only be designated with explicit support from the local community. They are established as areas of high nature conservation value with demonstrably enthusiastic local communities interested in promoting the sustainable development agenda. This network is being used and evaluated within COBWEB to assist in developing, testing and validating our concept of a citizen observatory. COBWEB utilises Biosphere Reserves in Wales, Germany and Greece to facilitate comparison of different aspects of the infrastructure platform across Europe. This short paper outlines the generic infrastructure platform solution as developed and demonstrated within these Biosphere Reserves.

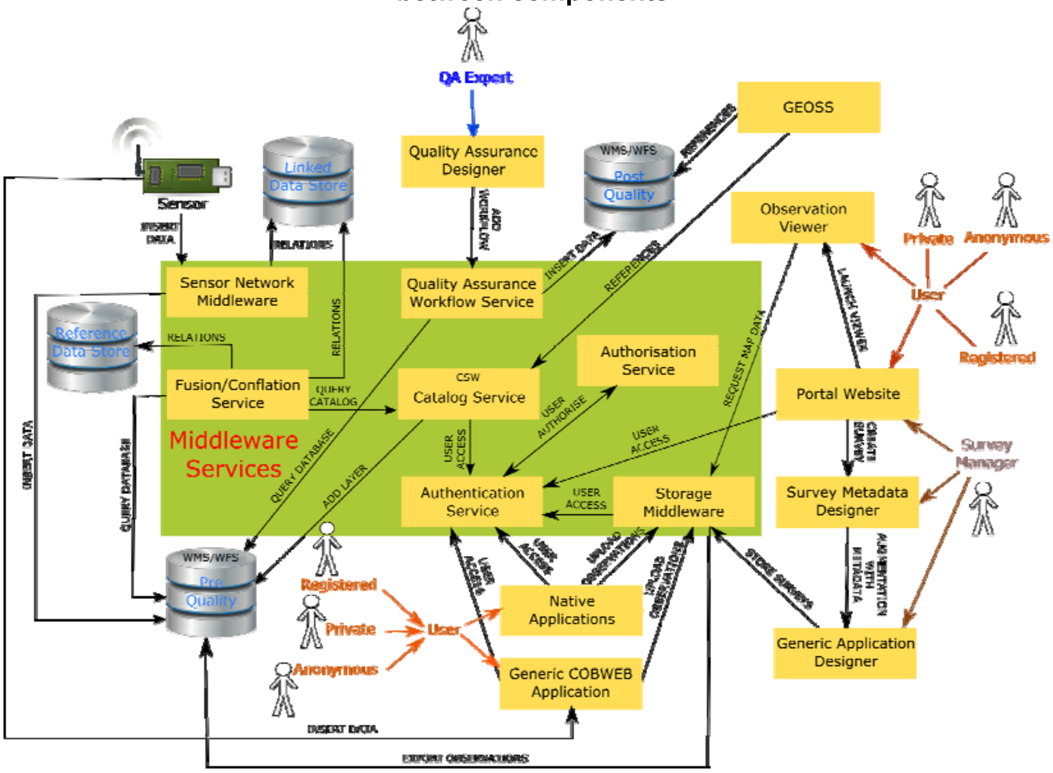
SYSTEM DESIGN THROUGH STAKEHOLDER ENGAGEMENT AND CO-DESIGN

Requirements for system design were initially gathered through a process of structured interviews. Starting with citizen groups located within the Biosphere Reserve study areas, the ideas were refined further through wider stakeholder engagement; in particular with groups associated with environmental governance and stewardship. To further assess the viability of the proposed solution, and to better understand the needs of the citizen, COBWEB engaged in a period of structured co-design activity. The use of regular workshops, interviews, and feedback from fieldtrips facilitated a deeper understanding throughout this activity. This allowed the realisation of real world requirements guiding the system design, based on real user needs ISO (2010). The co-design activity resulted in seven volunteer groups mobilising citizen scientists within the UNESCO test-bed areas, throughout the 2015 field season. This allowed the project to demonstrate, validate and improve the concept of a Citizen Observatory.

ARCHITECTURE

The COBWEB system architecture (Figure 1) has developed through a combination of new software builds, resulting from periods of rapid prototyping, building upon open source software. Whilst continuing to follow the requirements derived through stakeholder engagement and co-design activites. The architecture consists of the following key components: portal website, generic application designer, apps, storage middleware, quality assurance and conflation, sensor networks, whilst also implementing access control and privacy, and open standards. Together offering the ability to deliver the required generic infrastructure platform which facilitates the collection of citizen science data for the purpose of environmental monitoring and governance.

Figure 1: An overview of the COBWEB architecture, and the high level interactions between components



1.1. Portal Website

Currently integrated with the Dyfi Biosphere Reserve website for demonstration purposes, this is the main point of entry to COBWEB instances. At the centre of the portal is the COBWEB version of GeoNetwork; an implementation of the Open Geospatial Consortium's (OGC) Catalogue Services for the Web (CSW) standard. This version of GeoNetwork supports the concept of 'surveys', 'fieldsessions' and associated metadata.

Users can login and request to join selected citizen science communities running surveys whose participants can either be anonymous, private or registered. Once users have contributed observations via their mobile device, results are available for visualisation via the portal.

1.2. Generic Application Designer

For each community, there is at least one ‘Survey Manager’ whose privileges entitle them to setup and create surveys. COBWEB employs a hybrid App approach enabling survey managers to build custom data collection forms at the portal website, which can then be synchronised with the generic COBWEB application on individual users mobile devices (Butchart 2013). A wide variety of form elements are available to cater for a broad range of user requirements.

1.3. Apps

The generic App solutions offers the citizen the abilty to install onto their Android device, and login using various identity providers. The citizen then has the ability to either contribute to a publically available survey or a restricted survey which they have been given access to by the Survey Manager. The citizen will then be presented with the form designed by the Survey Manager, allowing them to participate in data collection with or without network coverage. In addition to this, the generic capability described in Section 1.2 is complemented by functionality allowing the cacheing of high quality basemapping on individual handsets for use in areas of poor or no network coverage.

To demonstrate the effective ‘separation of concerns’ in the architecture, and how the COBWEB framework can be used in scenarios where lower level access to inbuilt mobile device functionality is required, a native Application in the flooding thematic area case study area was also developed. This uses the same interface as the generic App for communicating with the ‘Storage Middleware’. (Figure 1).

1.4. Storage Middleware

Storage Middleware recieves the observations from the App. As long as Oauth v2 authorisation is supported, the Storage Middleware component provides a generic REST-based API accessed storage compatibility layer on top of a range of cloud based providers (Google Drive, Dropbox, etc.) or physical storage media where local storage is required.

Storage Middleware is a central component of the COBWEB architecture used for managing survey schemas and exporting geospatial observations to the desired encodings, e.g. KML, Geopackage, Shapefile, GeoJSON, CSV, etc. By synchronising all stored information with a relatonal database (the pre-quality PostGIS database in Figure 1) export of data via OGC Web Services (WMS/WFS) is supported.

1.5. Quality Assurance and Conflation

This is an important research area as a frequent observation made of citizen science sourced data is that, while there are large volumes of data, there quality is unknown making them of limited use. Within COBWEB, the approach has been to research how a variety of data provided by formal and informal crowdsourcing activities, observations from the co-design projects, sensor feeds and social media, could be used to achieve a measure of quality that could be expressed in the metadata.

COBWEB has designed a prototype for quality assurance (QA) using a standards based web service chaining approach, allowing great flexibility in what quality processes are applied to the citizen science captured data. This generic capability of authoring the QA is necessary as what quality control processes are applied is highly use case dependent. Besides allowing to reuse easily other web services within the QA workflow, specific quality controls generating quality elements are broadly grouped into seven Pillars that extends previous typology of quality assessment types (Goodchild and Li 2012): 1) Location based services, 2) Cleaning, 3) Automatic validation, 4) Comparison with authoritative data, 5) Model based validation, 6) Big/Linked data, and 7) Semantic harmonisation (Meek et al. 2014, Leibovici et al. 2015b).

The solution developed is based on the OGC's Web Processing Service standard and OMG's Business Process Markup Notation. Atomic quality controls are encapsulated as WPS processes that are composed and orchestrated using a workflow environment. The JBPM suite (workflow editor and workflow engine) have been customised to work with OGC services (Meek et al. 2015) then integrated with the portal website. The Survey Manager has the authority to create these quality assurance workflows.

Data fusion and Data conflation in COBWEB is used either after for final data use or during quality assurance for some validation of observations with external resources available on the Web (Wiemann et al. 2015, Leibovici et al. 2015). The process consists of a number of sub-processes, including data search and retrieval, data enhancement and harmonization, similarity measurement, data matching, evaluation and resolving (Wiemann and Bernard 2015) and are also accessible via WPS interfaces. Provenance is registered as within a Linked Data store that allows to finalise the data fusion data conflation.

1.6. Sensor Networks

Though data collection from mobile devices is fundamental to COBWEB, a variety of different sensors platforms, monitoring multiple environmental parameters within the Dyfi Biosphere Reserve testbed, has been deployed. These sensor networks have been specified in accordance to feedback from the co-design activities. In certain cases, sensor readings can be garnered by mobile devices and fused with observations. In other cases, data are routed to back-end servers where a conflation process can be initiated. These data can then be exposed to one or more of the Quality Assurance Pillars for subsequent analysis.

In addition to physical sensors, data from virtual sensor feeds is also captured. Such feeds are usually captured from pre-existing sensor network configurations; in the case of COBWEB, a legacy hydrological network has been harnessed. In this way, external data sources to COBWEB can be integrated into an arbitrary survey when data are made available in a public and standards-compliant way. Going forward, it is envisaged that public authorities will increasingly make selected data sources available in this way.

1.7. Access Control and Privacy

Initial stakeholder engagement revealed a requirement to be able to control access to sensitive data; for example, species protected under the UK Wildlife and Countryside Act. It is not desirable or permitted to make available over the web detailed information on these species. Conversely, these are often the most valuable data for consideration in environmental monitoring, and the type of information required for management and policy purposes.

In addition to data security, questions of privacy were also a requirement within COBWEB. It is essential to enable users to register using personal information so that decisions concerning what they are authorised to access and contribute towards can be made. Identity information has the potential to also be used for quality assurance purposes.

Exploiting previous work by Higgins 2012, COBWEB has further developed the use of an access management federation approach for securely sharing identity information. Based upon the OASIS Security Assertion Markup Language (SAML) and eXtensible Access Control Markup Language (XACML) standards, the key advantages of this approach are that it is a proven, industry strength solution that allows Single Sign On to protected web based resources across administrative domains. This means that citizens can access both protected and unprotected data sources, and collect and share protected data with public authorities in compliance with Data Protection legislation. Public authorities can leverage the benefits of interoperability, for example, with OGC web services, and potentially access all citizen sourced data without recourse to mechanisms such as anonymisation, obfuscation, reducing resolution, etc.

1.8. Standards and spatial data infrastructures

COBWEB has a requirement to make data collected through the infrastructure available within the Global Earth Observation System of Systems (GEOSS) without restriction. This has been addressed using a cooperative approach with the broader geospatial/citizen science community, with the development of a profile of the relevant OGC standards to maximise interoperability (swe4citizenscience 2015).

This has resulted in progressing a vision of a harmonised common data model, to which data can be published using OGC web services. Once realised, and with sufficient community support, most, if not all, crowdsourced, citizen science type data can be published to this open standard. This will then increase the immediate usefulness of these data and allow the myriad of potential users of such data to exploit existing standards based tooling and develop new standards based solutions. Integration costs will be reduced.

SUMMARY

COBWEB has developed a common framework for mobile device apps for use in citizen science for environmental monitoring. It is has shown that creating a generic solution to automating quality control and assurance which is sufficiently flexible to address the huge range of potential scenarios is beneficial to the reuse of citizen science data. Further development would result in the ability to make very large volumes of data useable, and is an area of potential future research.

Similarly, further attention needs to be paid to addressing whether it is possible to create a useable framework which is sufficiently flexible to allow a broad range of different kinds of familiar semantic resources to be employed in designing surveys before citizens go into the field. Without this, despite post-processing server-side, continued problems associated with a lack of semantic interoperability may be anticipated.

Despite perceived complexity and proliferation, the use of open interoperability standards still presents the most realistic chance of preventing the waste of resources and reuse opportunities inherent in creating silos of data locked into proprietary solutions. Standardisation effort should continue and adherence be required to help realise investment in spatial data infrastructure type initiatives such as GEOSS.

ACKNOWLEDGEMENTS

This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 308513.

REFERENCES

Butchart, B., Pope, A., King, M., Hamilton, G., Terzis, P. and M. Koutroumpas (2013). Fieldtrip GB: Creating a customisable mapping and data capture app for the HEFE community, at http://gogeo.blogs.edina.ac.uk/files/2013/04/gisruk2013_submission_85.pdf, [accessed 15 October 2015]

Goodchild, M.F. & Li, L. (2012) Assuring the quality of volunteered geographic information. Spatial Statistics. [Online] 1 (0), 110–120. Available from: doi:10.1016/j.spasta.2012.03.002 [Accessed: 26 June 2012].

Higgins, C.I., Koutroumpas, M., Matheus, A., Seales, A. (2012) Shibboleth Access Management Federations as an Organisational Model for SDI, International Journal of Spatial Data Infrastructures Research, 2012, Vol.7, 107-124.

ISO (2010) Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems, ISO 9241-210:2010

Leibovici, D.G., Evans, B. Hodges, C. Wiemann, S. Meek, S. Rosser, J. and Jackson, M. 2015 On Data Quality Assurance and Conflation Entanglement in Crowdsourcing for Environmental Studies. International Symposium on Spatial Data Quality (ISSDQ 2015), 29-30 September 2015, La Grande-Motte, France, 4 pages abstract.

Leibovici, D.G. Meek, S. Rosser, J. and Jackson, M. (2015b) DQ in the citizen science project COBWEB: extending the standards. Data Quality DWG, OGC/TC Nottingham, September 2015, U.K.

Liu, Hai-Ying and Kobernus, Mike and Broday, David and Bartonova, Alena (2014), A conceptual approach to a citizens' observatory - supporting community-based environmental governance. Environmental Health 2014, 13:107

Meek, S., Jackson, M. and Leibovici, D.G. 2014 A flexible framework for assessing the quality of crowdsourced data. 17th AGILE conference on Geographic Information Science, 3-6 June 2014, Castellon, Spain

Meek, S Jackson, M and Leibovici, DG 2015 Addressing the quality assurance challenge for location-based crowd-sourced data through workflow composition of OGC web services. Computers & Geosciences (submitted)

swe4citizenscience (2015) swe4citizenscience at https://github.com/opengeospatial/swe4citizenscience [Accessed 10 October 2015]

Wiemann, S Meek, S M Chapman, C. Leibovici, DG Jackson, and Lars, B 2015 Service-based combination of quality assurance and fusion processes for the validation of crowdsourced observations. AGILE 2015, 7-9 June Lisbon, Portugal

Wiemann, S. & Bernard, L. (2015) Spatial data fusion in Spatial Data Infrastructures using Linked Data. International Journal of Geographical Information Science. [Online] 0 (0), 1–24. Available from: doi:10.1080/13658816.2015.1084420 [Accessed: 2 October 2015].

Citclops Data Explorer: exploring water quality in the Wadden Sea*

Luigi Ceccaroni[1], Filip Velickovski[2], Meinte Blaas[3], Marcel R. Wernand[4], Anouk Blauw[5] and Laia Subirats[6]

[1] 1000001 Labs, Spain, luigi@1000001labs.org; Eurecat, Spain, luigi.ceccaroni@ce.eurecat.org

[2] Eurecat, Spain, filip.velickovski@eurecat.org

[3] Deltares, The Netherlands, Meinte.Blaas@deltares.nl

[4] Royal Netherlands Institute for Sea Research, The Netherlands, Marcel.Wernand@nioz.nl

[5] Deltares, The Netherlands, Anouk.Blauw@deltares.nl

[6] Eurecat, Spain, laia.subirats@eurecat.org

Abstract

This paper describes software development and adaptation, carried out within the European project Citclops, to: interface different technologies; integrate different components; implement feedback from end-users as delivered via citizen-science interfaces; interpret the data collected; and deliver new information to the citizens.

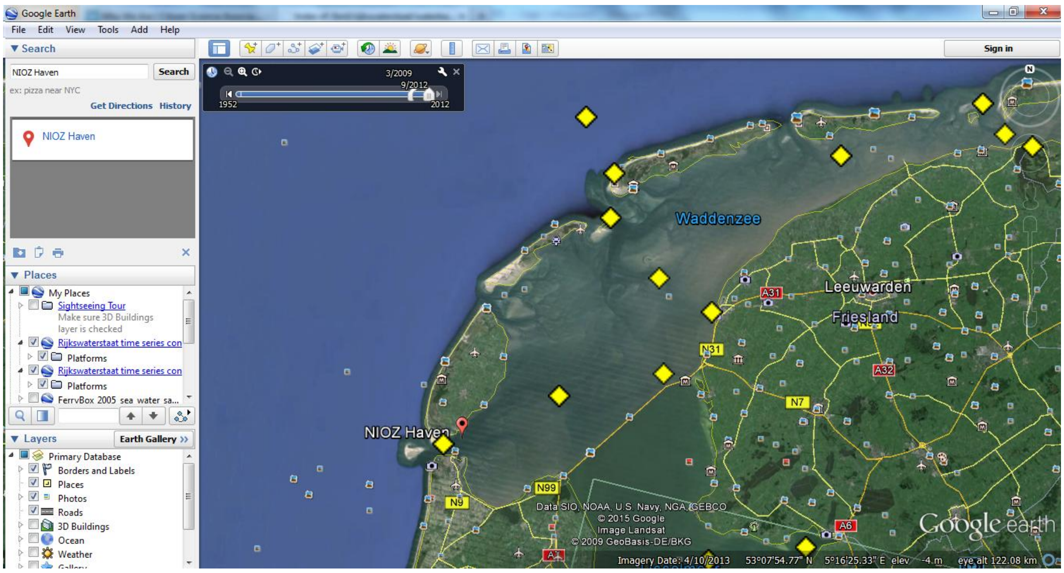
Keywords: citizen science, decision support system, data interpretation, knowledge-based systems integration, semi-supervised learning

---

1. INTRODUCTION

The Wadden Sea is a large-scale, intertidal marine system, where natural processes related to sediment transport and primary production define the basis of its internationally recognized cultural and ecological values. On the other hand, human pressures on the system abound: aquaculture, fishing, tourism and recreation, next to agriculture, mining activities and industry in the surrounding region. Nutrient inputs from rivers, growing tourism and large-scale fishery activities may create conditions that negatively affect human society, water quality and ecological systems. Satellites and in-situ monitoring are routinely used to collect information about water quality (see Figure 1). Recently, smartphone-based tools and other citizen-science sensors have entered the arena to enable also citizens to collect scientifically-relevant data (Graham et al, 2011). This paper describes part of the Citclops Data Explorer: a system to predict optical water-quality indicators in the Wadden Sea, e.g. for aquaculture, tourism and diving, but also for water managers. As *information sources*, the system uses MERIS satellite data, data collected with the *Citclops – Citizen water monitoring* app (Wernand et al, 2012) and physical data: waves, currents, river inputs and weather data for the period 2003-2015 (MERIS data available up to 2011; app data available in 2014-2015).

Figure 1: In-situ monitoring platforms



Source: Google

2. INDUCTIVE LEARNING

In this study, inputs form a **factored representation**, a vector of attribute values, and output is a discrete value. **Inductive learning** is used to analyse the input data and provide predictions in terms of *water colour,* using the *Forel-Ule* (FU) scale, a comparative scale firstly developed in the 19th century. (The FU scale has an implicit relation to other water-quality properties such as turbidity, transparency, suspended particulate matter and chlorophyll (Wernand, 2011).) In this way, a (possibly incorrect) general function or rule is learnt from specific input-output pairs. This is a case of **semi-supervised learning**, in which the system observes some example input-output pairs and learns a function that maps from input to output, with the caveat that the labels of the labelled examples may not be true. In the citizen-science setting of Citclops, a system is built to guess water's colour from a photo. Labelled examples are gathered by people snapping pictures and being asked about the colour, via an app. That is supervised learning. But in reality some of the people intentionally provide a wrong value, to test the app, for fun, or for whatever other reason. There are noise and inaccuracies in the data, and to uncover them is an unsupervised learning problem involving images, self-reported colours, satellite-measured colours and true (unknown) colours. Thus, in predicting water quality, both noise and lack of labels create a continuum between supervised and unsupervised learning.

In the Citclops Data Explorer, once *the ability to predict water quality* is learnt, it can be useful in several applications. Next to the abovementioned short-term direct use of apps by citizens, it can help water managers in longer-term monitoring, system analysis and decision making about water use. It will provide information to assess the constraints and opportunities for sustainable use of the sea and coast, and also to guide risk analysis and response to early warnings. With the information sources mentioned above, an inductive *learning element* (decision trees) has been used to predict water colour in the following week. The design of the learning element takes into account three major issues: (1) which *attribute* is to be learned; (2) what *feedback* is available to learn this attribute; (3) what *representation* is used for the attribute.

The *attribute* to be learned is water colour. The *type of feedback* available for learning has determined the nature of the learning problem that the system faces: *semi-supervised* learning, which involves learning a *function* from examples of inputs and outputs. The system learns a function from observations of MERIS satellite data, citizen data and physical data to a discrete output (colour represented as FU). Finally, the *representation of the learned information*, propositional logic, determines how the learning algorithms work. The last major factor in the design of the learning system was the *availability of prior knowledge*. The system begins with no knowledge at all about what it is trying to learn. It has access only to the examples in the data series. In this study, an algorithm for deterministic semi-supervised learning is given as input the value of the unknown function for particular inputs and tries to recover or approximate the unknown function.

2.1. Learning decision trees

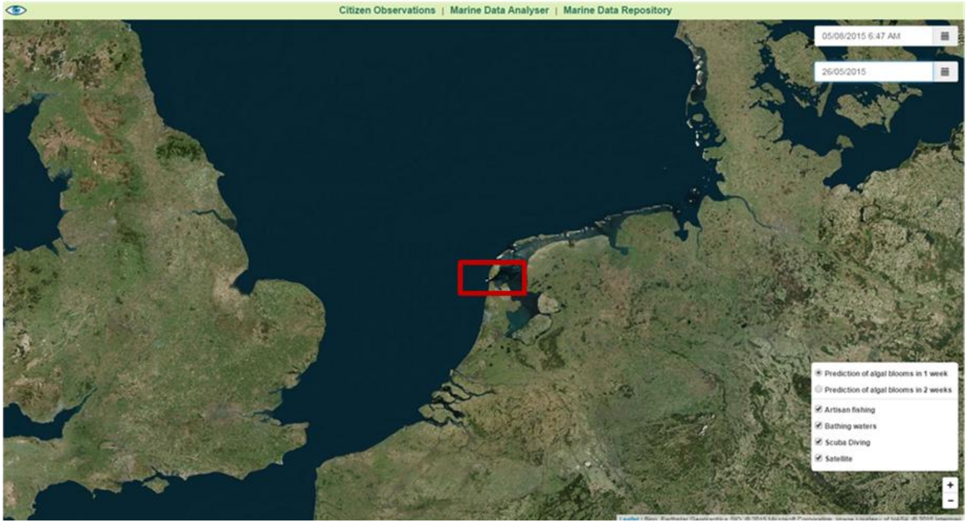
Decision tree induction is used in this study, being one of the most successful forms of learning algorithms and being a decision tree representation very natural for humans. The *decision trees* take as input a situation described by a set of attributes (from remote sensing, citizens and in-situ instruments) and return a *decision*: the predicted output value for the input, i.e., the prediction of the evolution of FU colour over a week's time. The input attributes are continuous. The output value is discrete; therefore this is a case of *classification* learning (the system is learning a discrete-valued function), wherein each example is classified according to the FU scale. The decision trees reach their decision by performing a sequence of tests. Each internal node in a tree corresponds to a test of the value of one of the attributes, and the branches from the node are labelled with the possible values of the test. Each leaf node in a tree specifies the value to be returned if that leaf is reached. The aim here is to learn a model for the **target label** *FU-Colour*.

2.2. Data description

We set this up as a learning problem and state what attributes are available to describe examples in the domain, which are the ones on the following list:

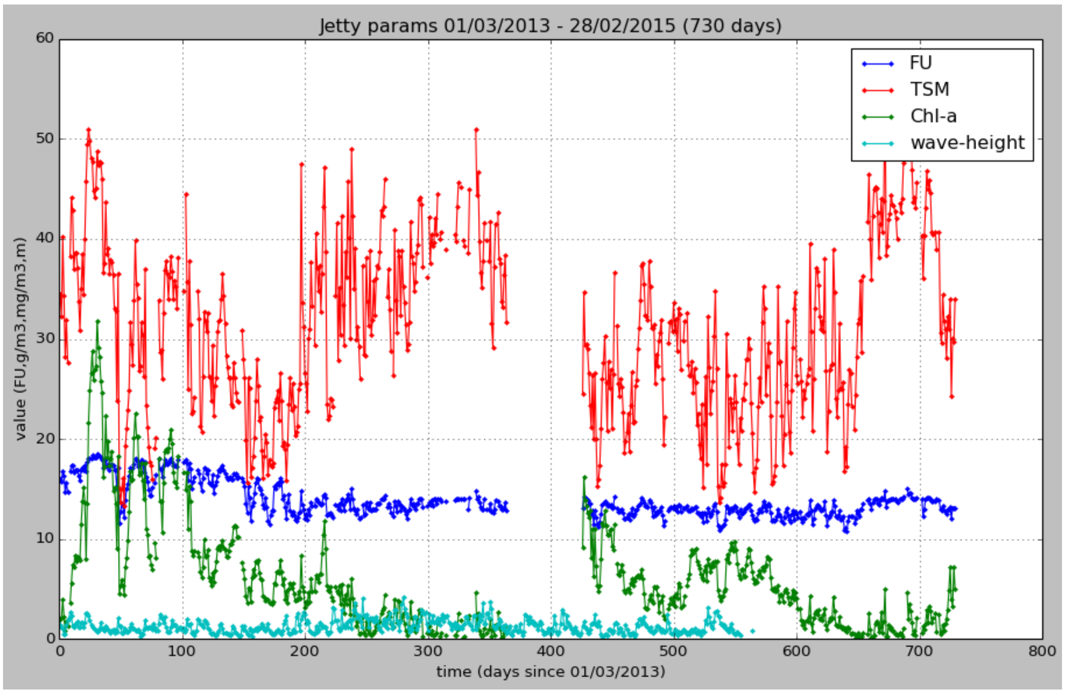
1. MERIS satellite data: FU, chlorophyll-a (2002-2011) - time resolution: one data point per day (missing data on cloudy days);
2. FU data collected with the *Citclops – Citizen water monitoring* app (2013-2015);
3. Water quality data: TSM, FU, chlorophyll-a collected in situ (2013-2015) (see Figure 2 and Figure 3) - time resolution: every two min during daylight (some missing periods);
4. Wave data (2003-2013) – average time resolution: one data point per hour;
5. Current data (2003-2013);
6. River inputs (2003-2013) – average time resolution: one data point per day;
7. Weather data (2003-2013);
8. SPM, chlorophyll-a, DOC, Kd collected in situ (2003 – 2013) – average time resolution: two data points per month.

Figure 2: Study area in the Wadden Sea for the 2013 – 2015 study period



Source: Citclops

Figure 3: Water quality data (TSM, FU, chlorophyll-a) and wave height (daily means)

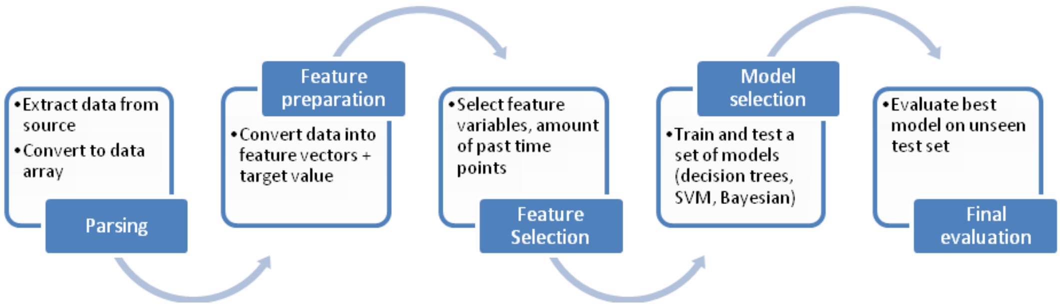


Source: NIOZ, Deltares and Citclops

3. METHODS

The following, available marine data have been assessed and used: **FU color** index, suspended particulate matter, TSM, dissolved organic carbon, light extinction, Secchi disk depth, chlorophyll-a, waves height, river input, weather data. A model of the target variable **FU color** at future points (2 days, 4 days, 7 days) has been learnt (see **Figure 4**). To do this, the initial problem has been converted to a three-class classification problem (see **Figure 5**).

Figure 4: Machine-learning pipeline



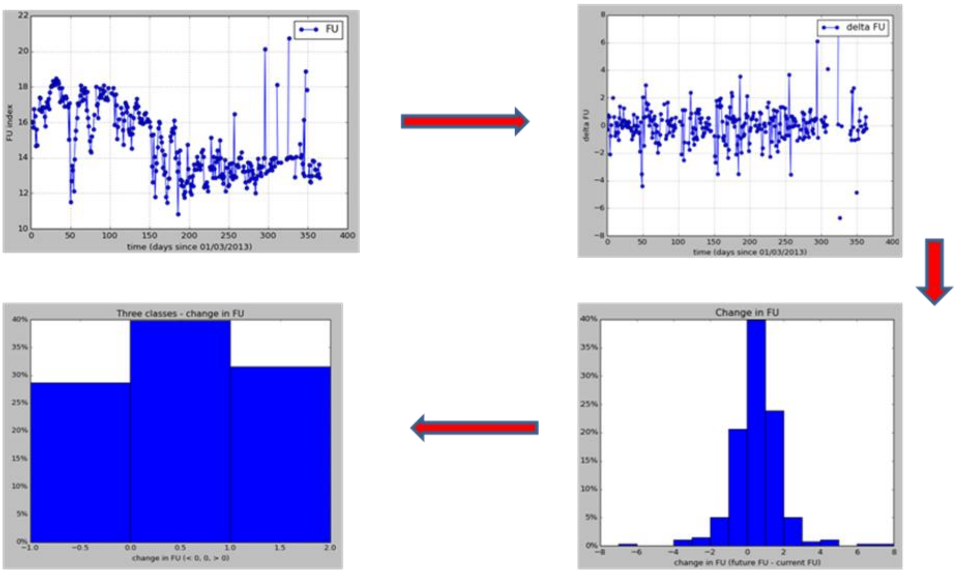
Source: Citclops

The model's prediction of FU has been evaluated using 10-fold cross-validation. The model has been integrated into the Citclops Data Explorer – Marine Data Analyser.

4. RESULTS AND DISCUSSION

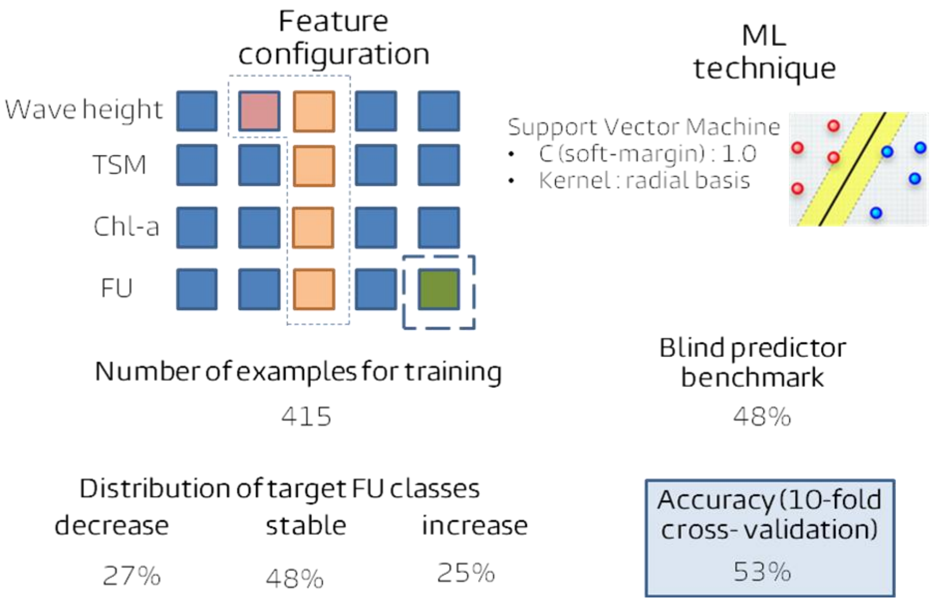
The task of finding optimal decision trees when the input descriptors are continuous is an intractable problem: time grows exponentially with the amount of data. Heuristics are therefore used to find solutions (deciding the sequence of tests and the specification of each test) in an acceptable time, but they are not guaranteed to be optimal. The forecasting system is composed of different decision trees, which predict the FU colour over a week, close to diving spots, fishing grounds and WFD/MSFD monitoring station. Results are presented in **Figure 6**, **Figure 7** and **Figure 8**.

Figure 5: Conversion to a three-class classification problem



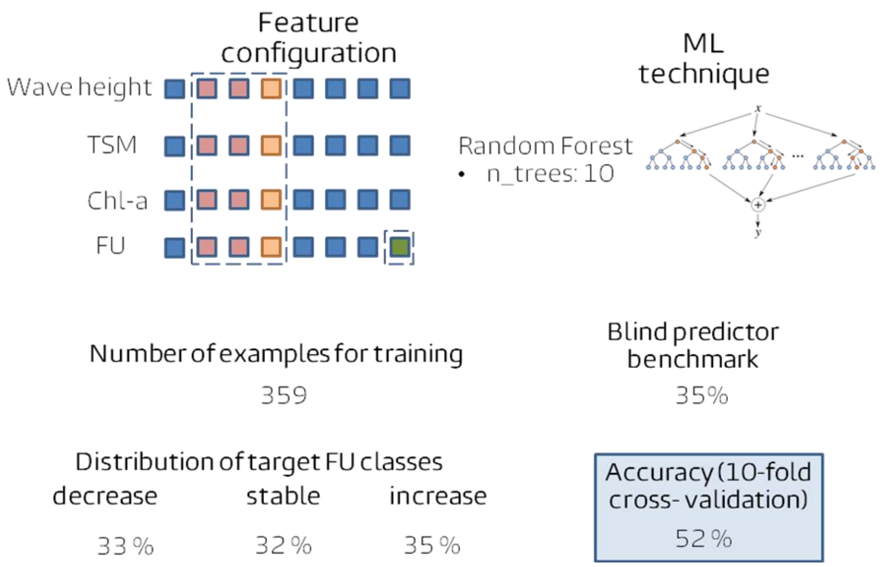
Source: Citclops

Figure 6: Example of results using a specific feature configuration and a support vector machine algorithm



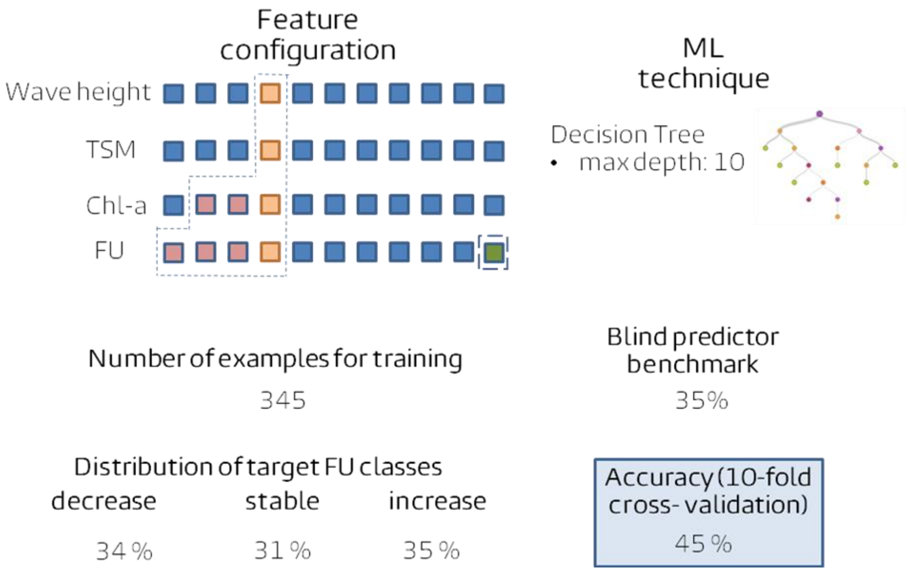
Source: Citclops

Figure 7: Example of results using a specific feature configuration and random-forest decision trees



Source: Citclops

Figure 8: Example of results using a specific feature configuration and decision trees with a maximum depth of ten



Source: Citclops

5. CONCLUSIONS

In this study, data from citizens, marine scientists and coastal planners are used to deliver information to citizens and policy makers, with the following potential applications: to simulate environmental crises; to chart emergency-management plans of the coastal zone; to provide sea farmers with bulletins about algal blooms; to maximize citizens' experience in activities in which water quality has a role; and to provide citizens with powerful, user-friendly tools of environment interpretation.

6. REFERENCES

Graham, E. A., Henderson, S. and Schloss, A. (2011). *Using mobile phones to engage citizen scientists in research*, EOS Transactions American Geophysical Union, 92 (38): 313–315.

Wernand, M. R. (2011). *Poseidon´s paintbox, Historical archives of ocean colour in global-change perspective*, PhD thesis, Utrecht University, pp. 240, ISSN 978-90-6464-509-9.

Wernand, M. R., Ceccaroni, L., Piera, J. and Zielinksi, O. (2012). *Crowdsourcing technologies for the monitoring of the color, transparency and fluorescence of the sea*, Proceedings of Ocean Optics XXI, Glasgow, Scotland, 8-12.

Citclops: data interpretation and knowledge-based systems integration*

Luigi Ceccaroni[1], Filip Velickovski[2], Alexander Steblin[3] and Laia Subirats[4]

[1] 1000001 Labs, Spain, luigi@1000001labs.org; Eurecat, Spain, luigi.ceccaroni@ce.eurecat.org

[2] Eurecat, Spain, filip.velickovski@eurecat.org

[3] Eurecat, Spain, alexander.steblin@eurecat.org

[4] Eurecat, Spain, laia.subirats@eurecat.org

Abstract

This paper describes the general interpretation of data and delivery of information as carried out via the development of a decision support system named 'Citclops Data Explorer' and available from the main portal of the European project Citclops.

Keywords: citizen science, decision support system, data interpretation, knowledge-based systems integration, Citclops Data Explorer

1. INTRODUCTION

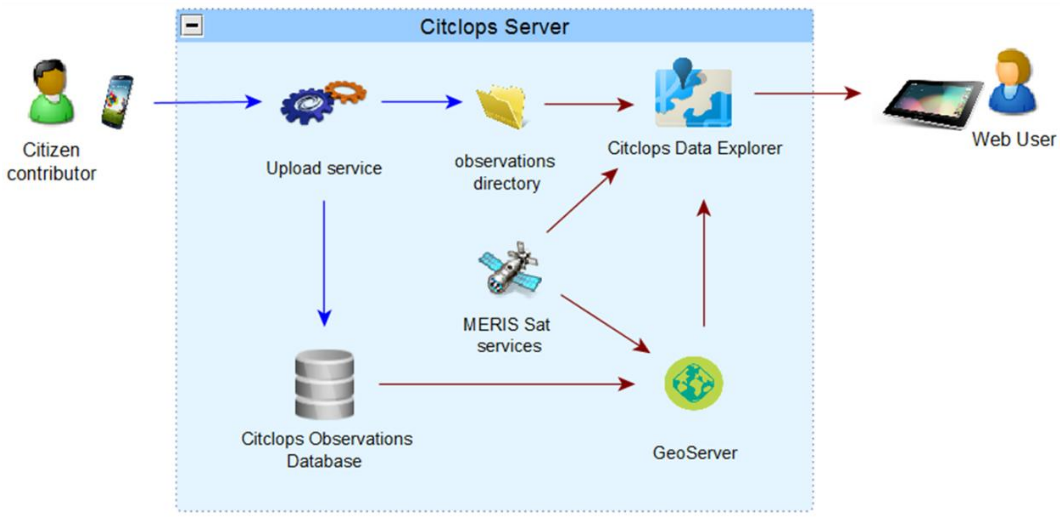
A decision support system named 'Citclops Data Explorer' has been developed and is available from the main portal of the European project Citclops, at [http://www.citclops.eu/participate/citclops-data-explorer]. The knowledge-based system behind the Citclops Data Explorer includes rules which relate sensor data streams, archived data sets and ecological status. EC directives that apply to the

---

domain of study of Citclops have been surveyed. The local context conditions have been also surveyed and taken into account via the integration of available results of regular monitoring by governmental organisations. The basic methodology that couples Citclops measurements to the ecological status (like colour to bloom phenomena) has been described.

Different platforms have been necessary in order to develop the components of the Citclops system (e.g., water-quality sensor technology, mobile and Web applications) and in this paper the interfaces defined among the different platforms and technologies are described. Of particular importance to the users has been the integration among the end-user apps for data collection, the user interface and the data server used as a repository. The final integration of all components into one single system can be accessed by the users via the project portal [http://www.citclops.eu] (see Figure 1).

Figure 1. Citclops's high-level architecture



Source: Citclops

2. REQUIREMENT SPECIFICATIONS

The requirement specification included the definition of use case scenarios, functionality of prototypes, performance evaluation and test cases. At the end of each project period, the prototype being developed has been published and evaluated. Data *quality-control* (QC) methods have been deployed during the project at various stages of development of data interpretation. The QC consists of a protocol with knowledge rules that can analyse the status of an observation of colour. Additionally, information is provided to the observer citizen via the app interface with directions to improve the measurements.

3. SPECIFICATION OF DEVELOPMENT PLATFORMS

Different technologies have been used and finally integrated into development platforms. A careful choice of the development platforms has reduced integration problems. A comparison of real-time systems for data processing has been carried out and resulted in the definition of the development platforms and technologies used in Citclops's back-end (see **Figure 2**) and front-end (see **Figure 3**).

Figure 2. Citclops's back-end technologies

Source: Citclops

Figure 3. Citclops's front-end technologies

Source: Citclops

4. DEFINITION OF INTERFACES

The communication infrastructure for the different components has been investigated and defined (see Figure 1). The interfaces among components of this communication infrastructure allow Citclops's architecture to be easily extendable so that future applications for monitoring and decision support can be added freely and without much effort. These component interfaces also allow for updates with new services or modified contextual elements.

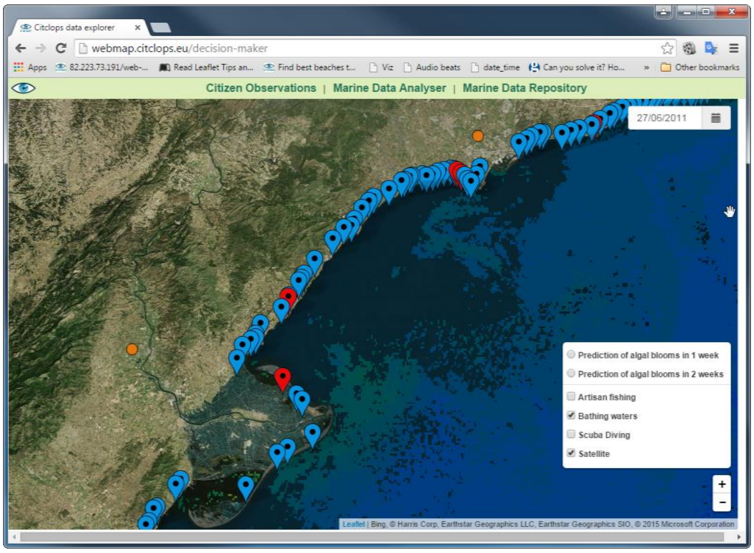
With respect to user interfaces, the Citclops Data Explorer presents three interfaces based on the content presented:

- Citizen Observations [http://webmap.citclops.eu/citizen-observations] (see Figure 4)
- Marine Data Analyser [http://webmap.citclops.eu/marine-data-analyser] (see Figure 5)
- Marine Data Repository [http://www.citclops.eu/search/welcome.php]

Figure 4. 'Citizen Observations' interface

Source: Citclops

Figure 5. The 'Marine Data Analyser' interface with the 'bathing waters' layer selected, showing the bathing waters and weather stations close to the Ebro Delta



Source: Citclops

4.1. Marine Data Analyser

The *Marine Data Analyser* (MDA) is a specific web map-based interface or view of the Citclops Data Explorer where an advanced user (such as a decision maker or marine researcher) can observe more in-depth information on the state of seawaters such as the historical state of bathing waters, seawater colour obtained from the MERIS satellite data and historic weather data. This view lets data collected from different sources be graphically visualised through time series plots in order to allow an analysis and reveal possible patterns or connections among variables.

The main map view of the MDA consists of optional layers (Bathing waters, Satellite...) that may be enabled or disabled in the bottom right corner of the screen. By default no layer is activated, and the view only shows a satellite map (provided by Microsoft Bing), on initialization centred at the geo-location of the user (if location is available and if user gives consent).

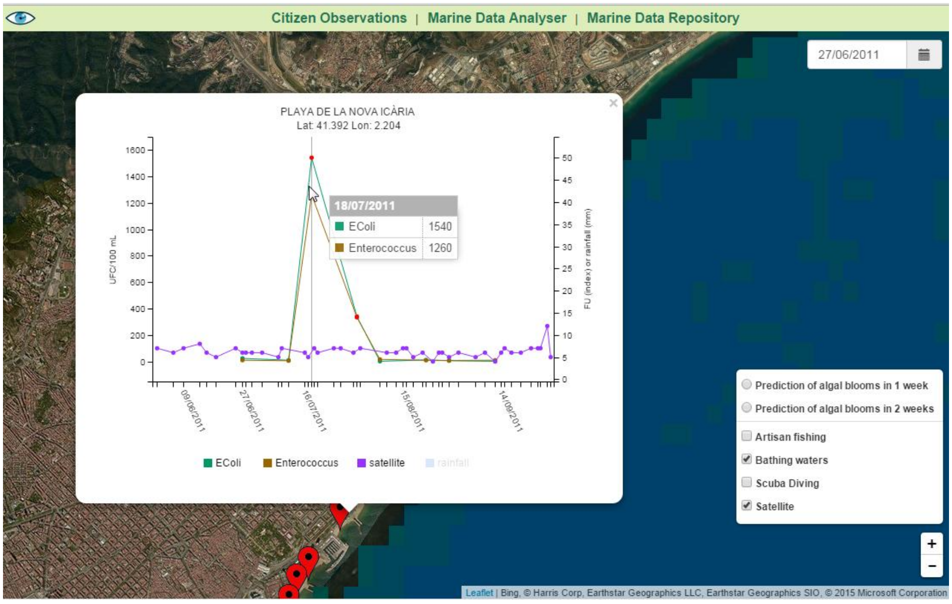
Since the 1970s, the EU has had rules in place to safeguard public health and clean bathing waters. The revised *Bathing Water Directive* (BWD) of 2006 updated and simplified these rules. It requires *members states* to monitor and assess their bathing waters for at least two parameters of faecal bacteria: Enterococcus, and E. coli (Directive 2006/7/EC). In the MDA, the summer bathing period (June – September) of the year 2011, and the bathing waters of the

Spanish autonomous communities of Catalonia, Valencian Community, and Basque Country were selected as prototype test zones and period, due to the availability of the data from multiple sources.

When the “Bathing waters” layer is enabled, as shown in **Figure 5**, the MDA interface places markers in geo-locations corresponding to beaches that have been officially designated as bathing waters. **Figure 5** is zoomed on the Ebro Delta coastal zone, showing bathing waters marked with blue and red markers. The blue markers represent indicators of normal levels of faecal bacteria as measured in the 2011 bathing period; red markers indicate that high or dangerous levels were measured. The orange circular markers are weather station locations, where data on rainfall were obtained.

When the “Bathing waters” layer is enabled, a graph of multiple parameters is displayed as shown in **Figure 6**. The user can enable and disable different marine data parameters of interest (E. Coli, Enterococcus, satellite FU colour, rainfall) and the interface will generate superimposed graphs, so that patterns may be observed and investigated. Additionally, specific data points are marked in red where the faecal bacteria have been recorded to have a level above the safe limit.

Figure 6. The Marine Data Analyser with the “Bathing waters” layer and Nova Icaria beach in Barcelona selected, showing time series of the following parameters: E. Coli, Enterococcus, satellite FU colour



Source: Citclops

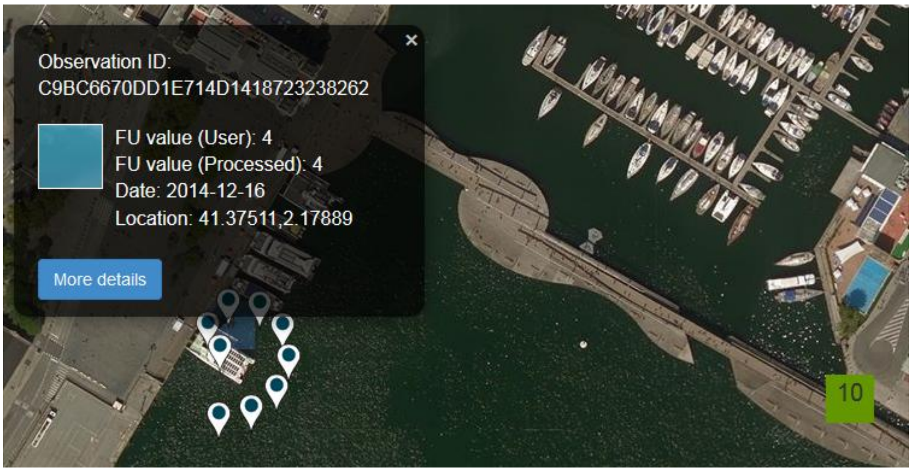
5. INTEGRATION INTO ONE PACKAGE

Sensor prototypes have been developed on different development platforms. Their integration consisted of making it possible to manage the sensors via apps. In this way all monitoring can be carried out by a mobile device. In this sense, all the different data-collection technologies have been integrated into one system and controlled via one interface.

With respect to information delivery, any citizen can navigate through all the observations of seawater, possibly grouped in clusters (see Figure 7), taken by user of the Citclops smartphone apps or uploaded manually using a dedicated web interface [http://www.citclops.eu/participate/upload-your-image]. Clustering polygons are defined by a radius of 80 pixels from the central marker and therefore are relative to the zoom level.

In addition, to ensure the quality of the colour samples, two quality-control **processes** have been implemented and integrated to filter and augment the citizen observation data. *Data quality control* (DQC) is an important aspect when dealing with samples collected within citizen-science initiatives; and the "Citclops - Citizen water colour monitoring" app allows any user to upload measurements without registrations, or any experience in the field.

Figure 7. Detail of the 'Citizen Observations' interface

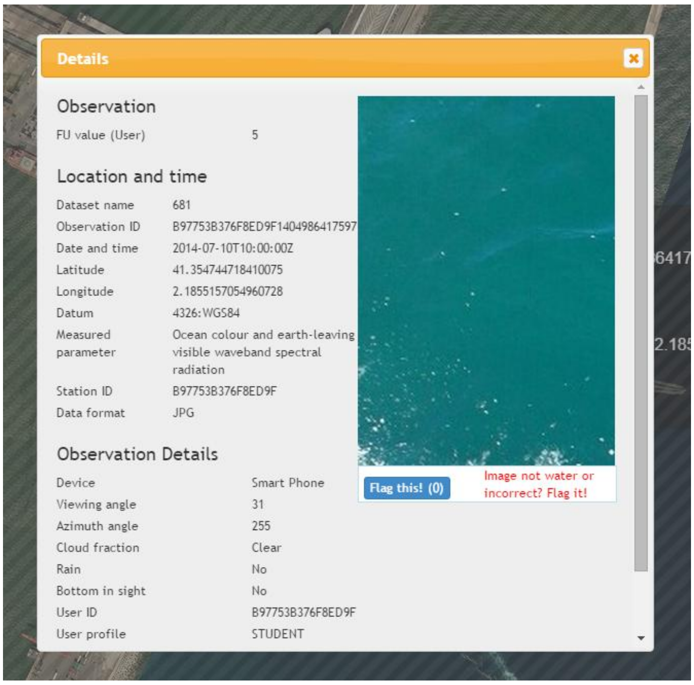


Source: Citclops

One of these processes is performed by a software module which computes the water colour from observation images collected via app. It works by locating, in the citizen supplied image, a sub-section that contains water, then computing the closest FU colour value of the sub section. (Both FU colour values are stored: the one indicated by the user via the app and the one automatically calculated by the software.)

Another process, to detect and remove invalid colour samples that do not correspond to a valid image, consists in the use of collaborative techniques, which have been implemented and integrated into the user interface. As can be seen in Figure 8, under the image there is a “Flag this!” button, which provides the possibility to a user to flag the image in case she thinks that the image is irrelevant within Citclops Data Explorer. Each user can vote only once per image. Between brackets, the number of flags accumulated by each image is shown. Once the image has been flagged, the Citclops Data Explorer administrator will receive an email and will decide whether or not to proceed with the removal of the image. Thus the collective collaboration of the users of the tool will contribute to maintain the quality of samples.

Figure 8: collaborative, quality-control technique in the Citclops Data Explorer



Source: Citclops

6. DEVELOPMENT OF CITCLOPS DATA EXPLORER

Raising environmental awareness and convincing citizens on the importance of environment measurements has been key to the success of Citclops. Web-based information is therefore presented in a way to attract a wide audience. In particular, the Marine Data Analyser (part of the Citclops Data Explorer) allows the deployment and execution of computer-based, multi-scale models related to retrieved data on seawater colour. Intelligent ICT tools have been deployed to predict the evolution of environmental variables and to present them via the graphical visualisation environment of the Marine Data Analyser, which also includes a modular and formal knowledge base and an inference engine.

As much as possible, the Citclops Data Explorer has been designed to be general enough to be applied to other domains. The Marine Data Analyser, for example, is a modular system designed to be applied to the prediction of any data represented with the same constraints as the data retrieved in the Citclops project. However, in practice, domain-specific knowledge is necessary, and the Citclops Data Explorer uses this knowledge for reasoning, to handle typically occurring cases in the relatively-narrow area of colour monitoring, in which the system has been validated (Ceccaroni et al. 2015). The user can access in-depth information on the state of seawater, such as the past state of bathing waters or the seawater colour obtained from satellite observations. These data are collected from different sources and analysed by data mining algorithms. The results are shown by graphs or specific marks on the map (see **Figure 6)**.

In general, all collected information on colour, transparency and fluorescence is converted to knowledge and maps on the ecological status of the waters. This knowledge is compared (close to the coast) to the status that is required by the Bathing Waters Directive. The results can be used to provide recommendations to the multiple sectors that make use of these waters, like local policy makers, tourists (regarding water quality and health) and fisheries.

7. CONCLUSIONS

A prototype of a decision support system, the Citclops Data Explorer, based on citizen science in the domain of water quality has been defined and implemented. As first step, an interface with a fixed set of interaction options has been developed. The interfaces between the sensors output, the Citclops Data Explorer and the interface module of mobile devices have been described. Using a modular structure, context awareness was added to the prototype. The context consists of time, environmental conditions, the user profile and the availability of other sources of information. The interfaces among the prototype modules have been explained. The prototype contributes a more flexible and citizen-based system to the monitoring community, which eases the access of end users (citizens, decision makers and researchers) to acquired information. Also, the Citclops Data Explorer allows establishing a tie-in of citizen-science information and satellite mapping to provide citizens and decision makers with: an overall GIS impression of an area; post processing of collected data (filtration and integration); derivation of information on the ecological status of an area; and production of key figures (maps and histograms) and their interpretation. Because of the relative availability of data in interoperable formats in the data server, the Citclops Data Explorer for water quality is mainly based on colour data only. Nevertheless, the Citclops Data Explorer has been designed to be general enough to be able to include data on transparency and fluorescence as soon as they are available in the data server.

8. REFERENCES

Ceccaroni, L., Blaas, M., Wernand, M. R., Velickovski, F., Blauw, A. and Subirats, L. (2015). *A decision support system for water quality in the Wadden Sea*, Proceedings of the 47th International Liege Colloquium on Ocean Dynamics. Liège, Belgium.

Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32006L0007, [accessed 1 October 2015].

"mO4Rivers" Web Mobile App

Getting river status data from Citizens

Vítor Almeida[1], Nuno Charneca[2], Yoram Rubin[3], Joaquim Pinto[1], Manuela Oliveira[1]

[1]Portuguese Association of Environmental Education, projetorios@aspea.org, Portugal

[2]National Laboratory for Civil Engineering, ncharneca@lnec.pt, Portugal

[3]WebH_2O, yoram.rubin@webh2o.net, USA

Abstract

In order to raise public awareness of the need for a sustainable management of water resources, the Water Framework Directive has established public participation as a requirement in the process of conception and implementation of River Basin Management Plans. *Projeto Rios* is a project that aims the promotion of citizen participation in the collection of environmental data that allows the identification and characterization of river ecosystems to ensure their monitoring and conservation. The "mO4Rivers" (myObservatory for Rivers) Web Mobile App was developed to support and facilitate the data submission process and to easily share information and knowledge obtained from data analytics processes. The groups of citizens are invited to adopt 500m river stretches and to interpret fauna, flora and to measure specific water quality parameters. This article describes the *Projeto Rios* methodology and the functionalities offered by "mO4Rivers" and shows how it can be an asset to promote citizen participation in the collection of environmental data related to river water bodies.

Keywords: *Projeto Rios*, Web Mobile App, citizen participation, river stretch

1. Introduction

Water bodies are among the most degraded ecosystems on the planet, suffering higher biodiversity reduction rates than those that occur in terrestrial ecosystems. The decline in water and habitat quality are often among the factors that impoverish its ecological integrity (Teixeira *et al.*, 2008). Achieve a sustainable management of watercourses will only be possible with a strong involvement of a more informed society and committed in the treatment of problematic change and degradation of water resources. The Water Framework Directive defines as fundamental the involvement of the population in the process of preparation and implementation of management plans of river basins in order to raise awareness, increase acceptance and commitment and promote a sense of belonging. (EU, 2000)

Citizen science is increasingly being recognized as an important new component of environmental monitoring and *Projeto Rios* (www.projectorios.org) is a project that aims to respond to the lack of an effective involvement of citizens in the problems concerning the deterioration of rivers water bodies ecological status.

The project results from *Projecte Rius,* launched in Catalonia by "Associació Habitat for Projecte RIUS Catalunya" in 1997. Since then has proved to be a success. In Spain, the project has more then 15 years of experience developing their volunteer activities, evolving so far more than 1000 groups in five autonomous communities (Catalonia[1], Galicia[2], Valencia[3], Cantabria[4], Madrid[5]). Through a protocol established between the Portuguese Association of Environmental Education (ASPEA[6]) and the "Associació Hàbitats for Projecte RIUS Catalunya", the project network was extended to the Portuguese territory in 2006 (Projeto Rios, 2014).

Using an experimental scientific method to collect and record environmental data, the groups of the *Projeto Rios* implement an adoption plan of a river stretch of 500 meters that include monitoring activities. Thus, the project pretends to promote a scientific curiosity and an affection for river ecosystems that would lead to a conscious change in the citizens behavior and, consequently, contribute to enrich the knowledge about the river bed and embankments and to eventually support decision making processes with the objective to improve the overall river ecologic status (Projeto Rios, 2014).

The key for a continued citizen involvement is the availability of detailed, relevant and understandable information (EEA, 2014). Therefore, the development of “mO4Rivers” Web Mobile App arises in order to support a simple and quick data submission process that conducts to an easy share of the information and the knowledge obtained from data analytics processes.

---

[1] www.projecterius.cat
[2] www.proxectorios.org
[3] www.xuquerviu.es
[4] www.proyectorioscantabria.org
[5] www.territoriosvivos.org/proyectorios
[6] www.aspea.org

2. *PROJETO RIOS* METHODOLOGY

2.1. Registration, selection of the river stretch to adopt and diagnostic field visit

All interested in being volunteers and actively participate in national network of *Projeto Rios* must fill an application form identifying the responsible for the group and the river stretch to adopt. All stakeholders can sign up and participate actively in *Projeto Rios*: schools, associations, private and public companies, municipalities, scouts groups, third age homes, NGO's, groups of friends and families. However, the groups should incorporate, in addition to a responsible, a support technician, a patron and the remaining participants, amounting to a minimum of four people. Currently, most of the registered groups are educational institutions, from kindergarten to university.

The first field visit marks the start of the group activity and, being the first contact with the river stretch, aims to take stock of its main dysfunctions and status, taking into account accessibility and security issues, the monitoring points where, in all future field trips, the monitoring measures will be carried out. During this first approach, the group fills the first field visit form (Figure 1) to systematize the knowledge and the needs assessment and thus prepare the subsequent field trips.

2.2. Detailed field visits and data collection

After a diagnostic, field visits start to take a much more detailed approach incorporating the analysis of physical-chemical, biological, hydromorphological parameters, biodiversity and data on land use, traditions and cultural and built heritage. Groups should make a minimum of two annual field visits (preferably in spring and autumn).

The project provides kits financed by the patrons, which contain files and tools to facilitate the interpretation of fauna, flora and field forms for recording the data of the river ecosystem.These field forms were, so far, available only on paper and did not allow, for that reason, the application of a mechanism to collect and centralise digital data generated by field trips.

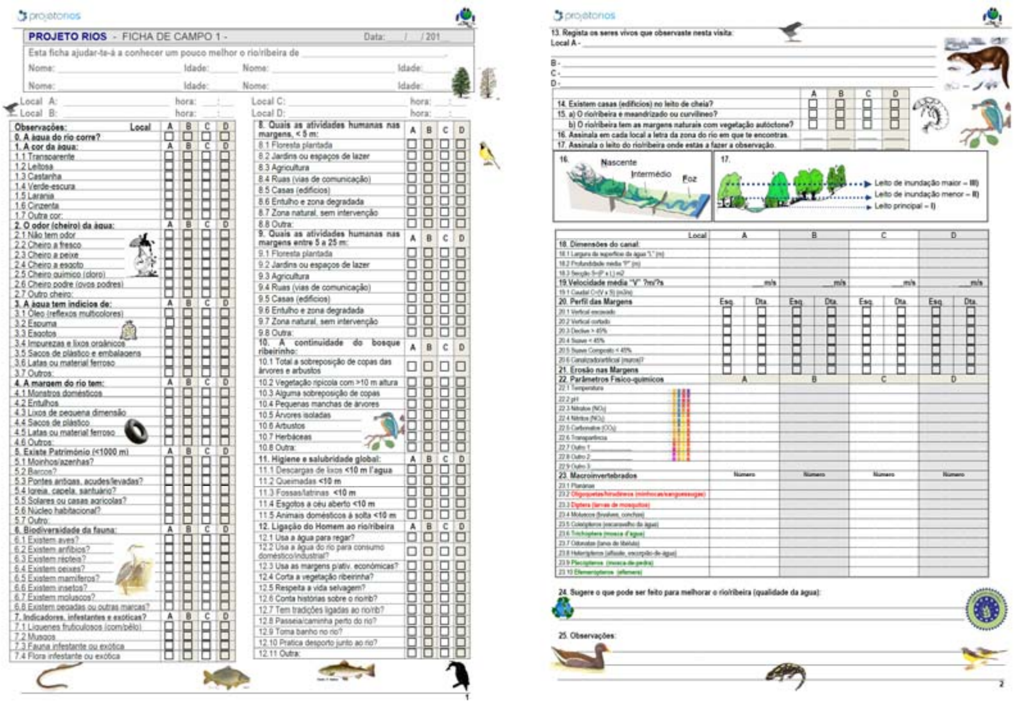

PROJETO RIOS - FICHA DE CAMPO 1 -

Figure 1 - Diagnostic field trip form

3. MO4RIVERS WEB MOBILE APP

The need to develop "mO4Rivers" arises from the difficulties to collect and submit the data and processing it. The centralisation of data using the App and myObservatory[7] features will allow the volunteers and managers of the project disseminating the knowledge acquired and to encourage the adoption of more rivers. This will also allow the use of these datasets by researchers, teachers and authorities in a easy, georeferenced and documented way.

3.1. Preparation of base data

The *Projeto Rios* methodology is based on the hydrographic network and on the identification of 500 meters segments (including the ones over reservoirs and estuaries). The geographic information available for Portugal (excluding islands) have about 120 000 kilometers long, which entailed an intensive editing work performed in a GIS environment that used the official hydrographic network provided by the Portuguese Environment Agency (APA). The resulting river stretches are identified in the spatial database implemented with a unique

---

[7] http://my-observatory.com/

identifier code (*RiverStrechID*), used to associate all the correspondent monitoring reports. The represented rivers were named based on the information present in the military official map (scale 1: 25,000). When and if a river segment is identified as missing the project managers should encourage the use of OpenStreetMap to vectorise the missing segments. This is also a way to disclose the features of this free geographic data project.

Both geographic and alphanumeric data of the project are registered in a PostgreSQL database with PostGIS extension. The implemented data structure contains about 50 tables and is used to: i) manage the registration process groups; ii) recording training; iii) support the answers to the field trip forms and diverse geographic information.

3.2. Application description

The "mO4Rivers" Web Mobile App[8] (Figure 2) is optimised to be used on mobile devices (smartphones and tablets) and in any PC browser. For georeferenced photos, notes and responses to interpretation questionnaires of river water bodies and associated ecosystems is recommended to have active GPS device. The application starts with the main menu which gives access to general functionality of the platform, divided into 3 groups (Record Data, View Data and Other):

a. *Take GPS–tagged photo*: allows the acquisition and loading of geo-referenced pictures in cases where the mobile device has the GPS location option activate. *Enter GPS-tagged notes*: allows, based on the GPS location option of the device, geo-reference notes.
b. *Trap Counts*: the count of individuals in traps for biodiversity characterization, usually insects, can be recorded on a suitable screen.
c. *My Forms Data Collection*: the application provides, for each of the available water system, questionnaires that reproduce as faithfully as possible the forms of *Projeto Rios* field trips.
d. *Collection Dashboard*: provides the overview of the answers to the questionnaires.
e. *View Data in Map*: the geo-viewer loads, as background map, the contents of OpenStreetMap (Mapnik) and, in the first access, focuses on continental Portugal. It offers two options: "Zoom To", which allows a zoom on the GPS location or on any other pre-defined geographic area, and "WMS Layers".
f. *View Data in List*: allows access to internal system data and its exploitation in terms of content.
g. *Field Event Calendar*: allows the management of various types of actions associated with the project.

---

[8] http://webh2o.net/mo4rivers

h. *Sync/Prep Offline Module*: In the case of lack of data connection in mobile equipment it is possible a later synchronization of the recorded data.
i. *Bulk Photo Upload*: allows the upload of multiple photos at once.

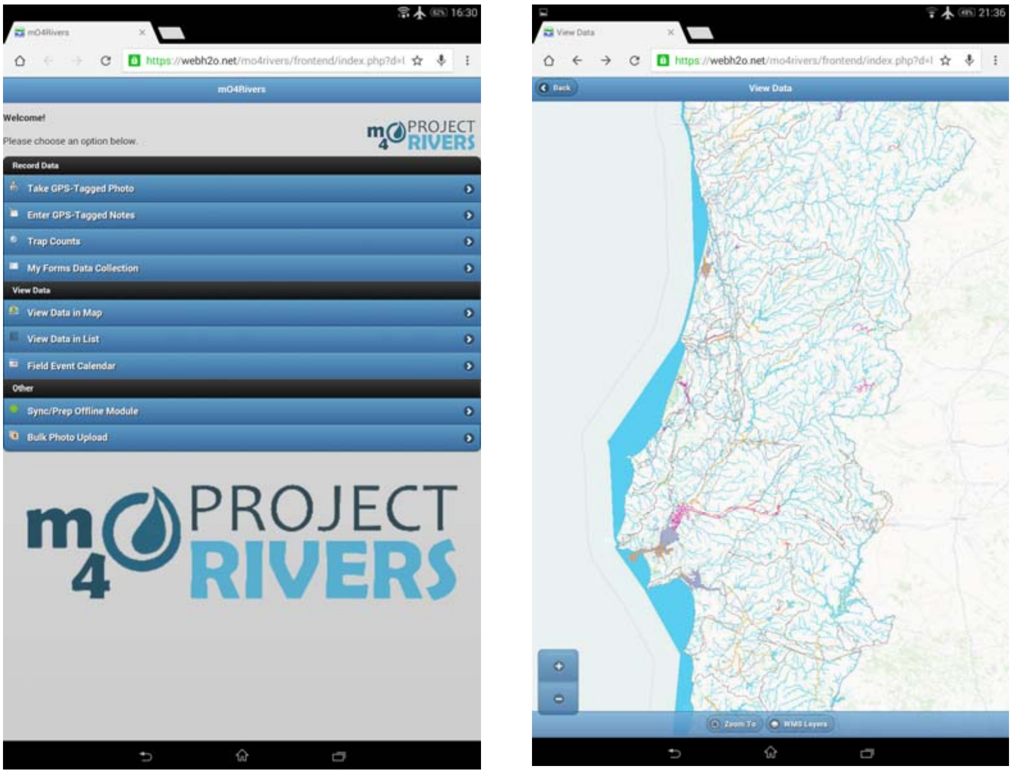


Figure 2 - “mO4Rivers” Web Mobile App screenshots

4. DISCUSSION AND CONCLUSIONS

This paper has briefly described *Projeto Rios* and the Web Mobile App “mO4Rivers” specific developed to support the data acquisition process by groups of citizens that joined the project through the adoption of 500m river stretches they visit, at least, 2 times per year. Using the application is possible to generate considerable amounts of *in-situ* data about river water bodies. The cost at which this data is obtained is far less than collecting authoritative data. In addition, the nature of crowdsourcing or citizen science means that information can be obtained over time allowing for new forms of systems evolution and because we are following an Open Access philosophy, anyone can use the information generated free of cost.

One of the most challenging aspects of this citizen science project includes overcoming the barriers to participation. The NGO supporting the project

(ASPEA) needs to do a continuous proximity work with all groups of citizens participating in the project, namely: schools, groups of scouts and other associations.

From the experience described above, several aspects should be taken into account, these include a wide adoption of the use of “mO4Rivers” App in all field activities of *Projeto Rios*. This adoption will provide, not only the support to identify and reporting of data collect in field trips, but also will promote a good communication with the community providing incentives for participation and the stimulation of scientific ideas. Moreover, developing for Android and IOS can be tricky as there are many different OS versions and phones/tablets available. The strategy for “mO4Rivers” App was to develop and deploy a web mobile solution, which doesn’t require an installation or update, directly from the user. The App is available through a URL and can be visited with a simple web browser.

Work will be done on the App to provide tips and helpful hints on how to do what without overburdening the users (and thereby losing their interest). The experience with the collection of data demonstrated the need for many easy-to-follow tips and hints. Also the timing of a campaign launch should be carefully chosen to match the target audience, e.g. if targeting schools, then campaigns should be integrated within the school year. Media launch events, i.e. press conferences, interviews and video for high-level media distribution, are important for initial and ongoing recruitment and outreach. Another key to success is to be well connected to grassroots organizations (e.g. the Associació Hàbitats in the case of the Iberian scale of the project). Outreach via social media is also an effective channel. Including a budget line for facebook-targeted campaigns that generate direct website clicks, website visits, etc. can yield additional participation.

In conclusion, engaging citizens in environmental monitoring via example given here has great potential to change the way we collect *in-situ* data and process it for public consumption. The project and the Web Mobile App that supports it offer the possibility to generate large amounts of timely, cost-efficient high quality information that was previously unavailable. The project team is know developing solutions to provide low-cost sensors kits to collect and transmit water quality data trough the Internet directly to myObservatory.

5. **REFERENCES**

EEA (2014). *Public participation: contributing to better water management. Experiences from eight case studies across Europe.* European Environment Agency. Luxembourg.

EU (2000). Directive 2000/60/CE (Water Framework Directive), 23 October 2000, European Union, Luxembourg.

Projeto Rios (2014). *Projeto Rios Presentation*. Portuguese Association for Environmental Education. Lisbon.

Teixeira, A., Geraldes A. M., Oliveira, J. M., Bochechas, J., Ferreira, M.T. (2008). Evaluation of Ecological Quality of Portuguese Rivers (AQUARIPORT Project): Summary of results for the analysis of benthic macro-invertebrate communities. IX Water Congress. Estoril Congress Centre, Cascais, 2-4 April 2008. pp. 1-12.

From Participatory Sensing to Making Sense

Mara Balestrini[1], Tomas Diez[1], Frank Kresin[2]

[1]Intel ICRI Cities, University College London, {m.balestrini, t.diez}@cs.ucl.ac.uk

[2]Waag Society, Amsterdam, frank@waag.org

Abstract

Crowdsensing or participatory sensing technologies are becoming increasingly available, allowing citizens to collectively monitor urban environmental factors. Initiatives that harness the potential of such tools are typically promoted by individuals, communities or organisations in a bid to foster bottom-up citizen participation in environmental action. However, our studies around diverse communities involved with the Smart Citizen project, an open source tool to produce and share environmental data, have revealed a number of challenges that hinder the uptake and sustainability of crowdsensing practices at the grassroots level. In this paper we summarise and discuss these findings and present Making Sense, a H2020 funded project aimed to overcome some of the challenges associated to the scalability and sustainability of environmental participatory sensing.

Keywords: Participatory sensing, environmental awareness, citizen participation

1. INTRODUCTION

Less than five years ago, most urban participatory sensing initiatives were part of research or citizen science projects, being deployed for relatively short periods of time and involving participants who had been recruited by the project instigators. In such experiences the focus was either on designing lightweight and reliable sensing technologies, and/or on creating the means to foster citizen participation in scientific research. More recently, with the proliferation of open source technology such as Arduino, the creation of makerspaces like Fab Labs, and the growing popularity of crowdfunding platforms, new urban sensing technologies have been designed and released to citizens without being part of specific citizen science projects or research agendas. Such novel technologies aim to empower self-organised groups of citizens by providing Do-it-yourself (DIY) or more open systems that they can appropriate for their own purposes. Notable examples are the Air Quality Egg (airqualityegg.com), Safecast (safecast.org) or RadiationWatch (radiation-watch.org).

There are some key differences between traditional citizen science projects and initiatives that emerge around open participatory sensing platforms. While in research and citizen science interventions the goals, engagement and operational strategies stem from the project instigators, who even train users or provide them with technology, in bottom-up initiatives the goals and strategies have to be negotiated by groups of citizens who gather around a concern or share a common purpose, need to gain access to technology and acquire the skills to operate them.

Although promising, the vision of truly bottom-up empowerment heralded by those who instigate open source participatory systems is hard to achieve. Over the past two years, our studies of user engagement with Smart Citizen, an open source participatory sensing platform, have revealed a number of technical and social issues that can hinder the appropriation of crowdsensing practices at the grassroots level. Lack of technical skills among users, difficulties with the usability and robustness of the sensing devices, a perceived lack of social interactions and purpose among community members, and problems with data reliability and meaningfulness have too often led to user disengagement with the platform. Other crowdfunded open projects faced similar issues. For example, in SafeCast the 10 most active volunteers have contributed almost 3/4 of the data, while Air Quality Egg has lost traction due to major technical shortcoming, issues around data ownership and reliability.

In this paper we discuss a number of themes that emerged from our previous research on user engagement with Smart Citizen and present Making Sense, a new H2020 funded European project that builds on them in the hope to make participatory sensing meaningful, sustainable and scalable. Making Sense will run three pilot interventions with the aim to co-produce the *Making Sense Toolkit*: a knowledge and resource online environment that provides access to technological resources such as software libraries, hardware firmware and digital designs for physical objects (for 3D printers); includes tools for data sensemaking; suggests strategies to sustain community engagement; and aggregates feeds from social media through which the pilot participants will share their collective and personal trajectories from digital making, participatory sensing and awareness to environmental action.

2. BACKGROUND

Researchers and practitioners, especially those investigating citizen science, have studied many aspects of participatory sensing systems, from technology design and features; to users' motivations to participate in citizen science and issues around data sensemaking (Bales et al., 2012; Canny et al., 2014; Corburn, 2005; Wiggins & Crowston, 2011; Willet et al., 2010). Most of these findings refer to either citizen science projects or research studies where participants have been recruited to collaborate with experts on agendas that have largely been defined a priorypriori. In consequence, there is little understanding of the factors associated with the appropriation of open source participatory tools at the grassroots level. How should such tools be designed in order to foster community appropriation? How can sustained engagement with crowdsensing be supported and scaled up? Furthermore, there's a lack of evidence on how these bottom-up technologies and practices might support individual and collective awareness, and ultimately enable communities to effect sustainable positive change. In the next section we present Smart Citizen and our studies around it.

2.1. Smart Citizen

Smart Citizen is an open source participatory sensing platform that comprises a sensor kit (SCK), an online platform and a mobile application. The project was launched in 2012, instigated by a group of stakeholders coordinated by the Fab Lab Barcelona (Diez & Posada, 2013). To date, over 1100 SCKs have been produced, of which 1022 are registered in the platform. Over 34 research institutions around the world have conducted research using Smart Citizen Kits.

Figure 1: The Smart Citizen data visualization platform (left) and SCK (right).

The SCK consists of an Arduino-based electronic board and shield, a battery, a Wi-Fi antenna, a MicroSD card, and a set of sensors to monitor humidity, temperature, nitrogen dioxide, carbon monoxide, sound, solar radiation, Wi-Fi hotspots, and battery charge level. The kit has been developed using open source technologies to allow advanced users to add features to their SCKs. The online platform (smartcitizen.me) allows users to upload data from their SCKs, share them through social networks and make them available to everyone online for free. Both the sensor kit and the online platform were developed with financial support from users through two crowdfunding campaigns.

In the last two years, we have conducted a number of ethnographic studies in order to assess user engagement with Smart Citizen (Balestrini el at., 2014b). We focused on three communities of users in Barcelona, Amsterdam, and Manchester, which had emerged in different ways. We collected data from over 100 users, through online surveys, semi-structured interviews, online postings and informal conversations. Following, we present and briefly discuss some of the common themes what emerged in all three interventions.

3. EMERGENT THEMES

From the 1100 SCKs that have been delivered to users around 624 have sent data to the online platform at least once. Only a small percentage of these sensors (around than 20%) are kept online, contributing data to the Smart Citizen platform in a sustained way. The following themes emerged as a result of our qualitative assessments.

3.1. Technological issues and lack of skills

Although SCKs are designed to be easy to set up, users usually lack the skills to install, operate and maintain sensing technologies. Even those who had reported to have "advanced technical skills", struggled to install their sensors. Our informants highlighted the need to have access to technical and methodological resources and assistance (e.g troubleshooting and documentation) that they could check online while setting up their devices.

To date, most low-cost sensors for environmental monitoring lack the robustness required to produce reliable data. In addition, users struggle to keep the sensors calibrated, which is crucial to obtaining reliable measures. This leads to random readings that have a negative impact on the quality of the data. Support for calibration can be provided both through an online platform and face to face technical meetups that, in turn, can foster social interactions and discussions leading to collective awareness.

3.2. Data sensemaking and meaningfulness

As in many platforms of the like, in Smart Citizen the shared data is represented online in the form of a flux or stream comprising lines and numbers. Most of our informants indicated that they struggled to make sense of such representations, and that this translated into a sense of "meaningless participation". We suggest that a sense of meaningfulness can be supported by adopting inclusive methodologies such as co-design (Balestrini et al. 2014), allowing citizens to collaboratively build tools and develop sensemaking techniques (e.g community displays, data annotation and comparison).

3.3. Lack of purpose

Our informants explained how they had lost interest in their sensors quickly after acquiring them. Even for those who had gone through the struggle of installing the SCKs to "just contribute some open data to a website" was not enough to sustain their engagement. They often highlighted the "need to find a purpose" to their contribution to participatory sensing. Unlike citizen science projects, self-organising communities need to negotiate common goals and shared purposes themselves. Platforms should provide social features for campaign organisation or methodologies to assist them in doing so.

3.4. Social interactions and championing

The Smart Citizen communities that we studied had emerged following different engagement strategies. While some were self-organised (the Barcelona community was formed by those who had crowdfunded the project and lived in that city), the one in Amsterdam emerged around an initiative hosted by the Waag Society (a cultural institution). Our findings demonstrated that those communities where members profited from social interactions, particularly in face to face settings, achieved larger levels of engagement and contributions. The role played by social interactions and community champions who can draw users to participate has been larlegy overlooked in previous grassroots digitally enabled communities.

The themes emerging from our studies around community engagement with Smart Citizen are aligned with findings reported in the citizen science and environmental justice literature: the success of a pilot intervention has been linked to factors such as providing training and skills through community coordinators, and following an approach that is context-specific, iterative, and adaptive (Pollock & Whitelaw, 2005). Furthermore, the fundamental role played by community champions to ensure the sustainability of community based environmental stewardship and civic action has been signposted in previous reports (Conrad & Daoust, 2008; Conrad & Hilchey, 2011, and Pollock & Whitelaw, 2005).

4. MAKING SENSE: A NEW APPROACH

Making Sense, which will officially start on December 2015, aims to foster and support the sustainability and scalability of environmental participatory sensing initiatives at the grassroots level. It builds on the hypothesis that open source software and hardware, digital maker practices and methodologies such as open design can be effectively used by local communities to appropriate their own technological sensing tools, make sense of their environments and address pressing environmental problems in air, water, soil and sound pollution. To achieve this goal, the project will develop a *Making Sense Toolkit* based on the existing Smart Citizen platform. The toolkit will be tested in three pilots in Amsterdam, Barcelona and Prishtina.

4.1. Consortia

The Making Sense consortium is interdisciplinary, including IAAC (Architecture and city planning), University of Dundee University (Duncan of Jordanstone College of Art and Design/DJCAD & Centre of Environmental Change and Human Resilience/CECHR), Waag Society (Institute for Arts, Science and Social Innovation), JRC (Foresight and Behavioural Insights Unit), the Peer Educators Network (Kosovo) and the European Fab Lab Network. Furthermore, Fab Lab Barcelona (part of IAAC) and Fab Lab Amsterdam (part of Waag Society) are involved in the project.

4.2. Approach

The project will provide actionable knowledge, existing tools (primarily Smart Citizen Kits) and networks (Fab Lab, maker spaces) as well as methods (Co-Inquiry, Three Horizons) to foster collective awareness on ecological related issues. The stakeholders will mobilise and support local communities to provide them with insight in their personal living environments.

The Making Sense toolkit will be co-developed, tested and iterated in three consecutive pilots in the cities of Amsterdam, Barcelona and Prishtina. To achieve these goals, the project will scope issues and build campaigns that matter to those who are involved in the pilots, to develop focused interventions and areas of interest, and to build and nurture communities around those. We will engage *communities of interest*, being groups of people who jointly perceive an environmental challenge in their local environments, and *communities of practice*, comprising hardware makers and tinkerers well versed in open source technologies and digital fabrication. We will create opportunities for them to meet at their local Fab Labs to co-design, test and improve readily available open hardware and software tools, and contribute knowledge about best practices around community-driven environmental sensing and data sensemaking. These creations will be shared across pilots and via the Making Sense toolkit for others to use and re-appropriate.

Furthermore, the communities will have opportunities to interact with experts and city officials, collect, share, visualise and interpret the data they collected, and devise ways to act on these insights, either individually or collectively.

4.3. Contribution

The project will develop a collaborative online Making Sense Toolkit as a knowledge and resource online environment that will foster community appropriation of existing sensing technology infrastructures, encourage data sensemaking processes, embed technical and methodological skills and foster learning among participants and across communities. It will be composed of an online repository that contains a database providing access to technological resources such as software libraries, hardware firmware (access through github.com) and digital designs for physical objects (suitable for 3D printers); integrates access to the Fab Lab.io, which is the knowledge and communication exchange supporting the Fab Lab network (developed by Fab Lab Barcelona); aggregates feeds from social media sites such as Instagram, Twitter, Youtube and Facebook through which users will share their collective and personal experiences during the pilots; and serves as the basis of environmental monitoring campaigns and citizen journalism initiatives to be conducted. The Making Sense Toolkit will be open to inspire and empower grass-roots environmental action all over the world.

5. CONCLUSION

Making Sense's goal is to pioneer a participatory approach that overcomes the many gaps that have hindered sustained and meaningful participation in previous urban participatory sensing interventions. Previous initiatives have failed to enable sustained and meaningful engagement mainly due to lack of data reliability provided by sensors, lack of technical skills among participants, lack of championing provided by local communities and experts, lack of feeling of purpose over the interventions. We bridge these gaps by connecting communities of makers with technical skills and communities of interest with drive and domain expertise. Moreover, we will orchestrate engagement with local institutions and community champions; we will foster a creative environment for purpose and ownership to emerge through the co-creation of bespoke solutions to tackle local challenges. Through these actions, we will develop a methodology that makes citizen sensing rewarding, sustainable and impactful.

REFERENCES

Bales, E., Nikzad, N., Quick, N., Ziftci, C., Patrick, K., & Griswold, W. (2012, May). Citisense: Mobile air quality sensing for individuals and communities Design and deployment of the Citisense mobile airquality system. In PervasiveHealth'12 (pp. 155-158). IEEE.

Balestrini, M., Bird, J., Marshall, P., Zaro, A., & Rogers, Y. (2014, April). Understanding sustained community engagement: a case study in heritage preservation in rural Argentina. In Proc. of CHI'14 (pp. 2675-2684). ACM.

Balestrini, M., Marshall, P., & Diez, T. (2014, September). Beyond boundaries: the home as city infrastructure for smart citizens. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (pp. 987-990). ACM.

Canny, J., Mascolo, C., & Paulos, E. (2014). Pervasive Analytics and Citizen Science. Perva- sive Analytics and Citizen Science, 13(2), 18-19.

Corburn, J. (2005). Street science: Community knowledge and environmental health justice. The MIT Press.

Conrad, C. T., & Daoust, T. (2008). Community-based monitoring frameworks: Increasing the effec- tiveness of environmental stewardship. Environmental Management, 41(3), 358-366.

Conrad, C. C., & Hilchey, K. G. (2011). A review of citizen science and community-based envi- ronmental monitoring: issues and opportunities. Environmental monitoring and assessment, 176(1-4), 273-291.

Diez, T., & Posada, A. (2013, February). The fab and the smart city: the use of machines and technology for the city production by its citizens. In *Proc. of TEI'13* (pp. 447-454). ACM.

Ganti, R. K., Ye, F., & Lei, H. (2011). Mobile crowdsensing: current state and future challenges. *Communications Magazine, IEEE*, *49*(11), 32-39.

Kera, D., Rod, J., & Peterova, R. (2013). Post-apocalyptic Citizenship and Humanitarian Hardware. *Nuclear Disaster at Fukushima Daiichi: Social, Political and Environmental Issues*, 97.

Pollock, R. M., & Whitelaw, G. S. (2005). Community-based monitoring in support of local sustainability. Local Environment, 10(3), 211-228.

Wiggins, A., & Crowston, K. (2011, January). From conservation to crowdsourcing: A typol- ogy of citizen science. In System Sciences (HICSS), 2011 44th Hawaii International Confer- ence on (pp. 1-10). IEEE.

Willett, W., Aoki, P., Kumar, N., Subramanian, S., & Woodruff, A. (2010). Common sense community: scaffolding mobile sensing and analysis for novice users. In Pervasive Computing (pp. 301-318). Springer Berlin Heidelberg.

Next-Generation Environmental Sensing: Moving Beyond Regulatory Benchmarks for Citizen-Gathered Data

Jennifer Gabrys[1] and Helen Pritchard[1]

[1]Citizen Sense, Goldsmiths, University of London

j.gabrys@gold.ac.uk

Abstract

A number of environmental sensing technologies and practices are emerging that seek to enable citizens to use DIY and low-tech monitoring tools to understand and act upon environmental problems such as air pollution. These "citizen sensing" projects intend to gather data sets, which can indicate environmental change and give rise to political action. Citizen-generated data sets are often gathered with equipment that diverges from state and regulatory standards, and through practices that differ from scientific protocols. When monitoring air pollution or other environmental disturbances with low-cost technology, citizen-led initiatives are often challenged about the validity or accuracy of their data. Environmental regulators often dismiss citizen-collected data as biased, lacking standardised procedures for collection, and having been generated through imprecise instruments. Yet environmental monitoring technologies are now deployed in multiple contexts, which might include not only consistently observing air pollution with sophisticated instruments to meet regulatory standards and to ensure compliance with air pollution policy; but also might include capturing fine-grained yet sporadic pollution data through DIY devices that provide an indication of changes in air quality, rather than an absolute measurement.

Keywords: Citizen sensing, citizen science, air pollution, participatory research, environmental regulation

1. INTRODUCTION

A number of environmental sensing technologies and practices are emerging that seek to enable citizens to use DIY and low-tech monitoring tools to understand and act upon environmental problems such as air pollution. These "citizen sensing" projects intend to gather data sets, which can indicate environmental change and give rise to political action. Citizen-generated data sets are often gathered with equipment that diverges from state and regulatory standards, and through practices that differ from scientific protocols. When monitoring air pollution or other environmental disturbances with low-cost technology, citizen-led initiatives are often challenged about the validity or accuracy of their data. Environmental regulators often dismiss citizen-collected data as biased, lacking standardised procedures for collection, and having been generated through imprecise instruments. Yet environmental monitoring technologies are now deployed in multiple contexts, which might include not only consistently observing air pollution with sophisticated instruments to meet regulatory standards and to ensure compliance with air pollution policy; but also might include capturing fine-grained yet sporadic pollution data through DIY devices that provide an indication of changes in air quality, rather than an absolute measurement.

As the U.S. Environmental Protection Agency has expressed in its analysis of new modes of environmental monitoring, "types of data" and "types of uses" are interlinked (US EPA 2013). Data typically only become admissible for legal claims when gathered through specified scientific procedures by state certified users with approved (as well as expensive) instrumentation. However, there may also be situations in which data gathered through citizen sensing practices are "just good enough" for establishing that a pollution event is happening. It therefore remains a relatively open question as to what the uses and effects of data gathered through citizen-sensing technologies might be, since these are data practices that are still emerging.

"Just good enough data" is then a phrase that surfaces when questions are asked about the accuracy of citizen-collected environmental data. "Just good enough data" draws attention to attempts to counter the reliance on high-levels of measurement accuracy as the sole criterion by which data are evaluated. What different practices emerge when environmental data are engaged with in a more indicative register? What do these practices enable? And what other relations, connections and points of focus might "just good enough data" generate?

Examining these questions in the context of participatory fieldwork conducted by the Citizen Sense research group in northeastern Pennsylvania, this paper considers how the use of air pollution monitors by residents living next to hydraulic fracturing (or fracking) infrastructure produces different registers and types of data. The paper outlines citizen-sensing practices that monitor fracking-related pollution that are already underway, and it discusses our attempts to contribute to monitoring processes through further participatory and practice-based citizen-sensing initiatives. The paper reviews the multiple forms of data generated through this participatory citizen-sensing project that that diverge from state and regulatory monitoring, including air quality data, data logs, citizen observations and stories, as well as a data analysis tool developed by Citizen Sense to facilitate citizen-led analysis of their data collected over 9 months of monitoring. The paper further discusses how residents attempt to mobilise data and engage in discussions with regulators, and the ways in which citizen-gathered data can provide other insights beyond a regulatory-only focus on monitoring.

Citizen-sensed data is rich with trans-local experiences and collective insights, often bringing attention to environmental change from new perspectives. Instead of reducing discussions of sensing practices to accuracy and regulatory alignment, how might we develop practices and infrastructures for a “just good enough” data that enables citizen-sensed data to make expanded contributions to environmental sensing? We suggest that the relevance of citizen-collected air quality data is not solely determined through absolute criteria, or alignment to state-managed air quality data, since these criteria can often shift depending upon modes of governance, location, and available resources. If data are understood instead as entities that transform depending upon the uses to which they will be put--and how “good enough” they might be to achieve these ends--it then becomes possible to attend to how data are differently mobilized through monitoring practices and political encounters.

2. AIR QUALITY MONITORING AND NATURAL GAS EXTRACTION

Unconventional natural gas extraction in the form of hydraulic fracturing began in the Marcellus Shale region in Pennsylvania in 2003, however by around 2006 the number of wells drilled in the state began to increase rapidly and communities began to notice more intensively the impact of the industry (State Impact). At the time of writing this paper, permits have been given for almost 17,000 wells, and just over 7,000 wells are in operation as sites of natural gas production, with more wells becoming active daily. As recorded by a local citizen-led website, MarcellusGas.Org, which collates and provides information and data on Marcellus gas well production in Pennsylvania, on average one new well was opened every two days in the state in September 2015.

Many of these wells and the related natural gas infrastructure of compressor stations, well pads, dehydrators, water impoundments, monitoring stations and pipelines are densely located in northeastern Pennsylvania. Along with this infrastructure, inevitable concerns have arisen about environmental impacts, especially in relation to water and air pollution. While much attention has been given to water pollution through several high-profile cases of contaminated well water in areas of northeastern Pennsylvania, residents of this community have also had concerns about the relatively under-monitored effects of fracking on air quality.

In order to understand the air pollution arising from the processes of natural gas extraction and production, residents of Pennsylvania have undertaken many diverse practices of monitoring with differing aims and objectives. In order to try to better understand their experience of air pollution from natural gas, residents have used an extensive range monitoring technologies in order to gain a more immediate sense of their environmental conditions. Many monitoring practices have required that residents collect samples for lab analysis that takes place in distant sites of data processing. Or that residents use technologies that produce data in forms that are not immediately useable. The promise of low-cost and next-generation environmental sensors is that data will be made available in real time to the users of the technologies.

In part, Pennsylvania residents' interest and sense of urgency in developing monitoring practices has also been in response to the lack of governmental monitoring in this rural area. Existing monitoring for the nationwide Air Quality Index (AQI), which is facilitated on a state level by the Pennsylvania Department of Environmental Protection (DEP), typically focuses more intensely on urban areas and roadside sites, and does not have a particular remit or attention to accounting for emissions from particular industries such as oil and gas. In this way, DEP stations for monitoring air quality and criteria pollutants such as particulate matter 2.5 (PM 2.5) are located in relatively distant urban centres such as Scranton, where monitors are often placed close to busy highways. Although the DEP also do undertake some mobile monitoring on a sporadic basis, due to economic and political constraints there is no consistent monitoring taking place by the regulators that would fully account for local emissions from the natural gas industry in the northeast of the state. Within the context of a newly expanding industry that residents felt was not sufficiently monitored, an interest then emerged to develop techniques for documenting environmental pollution in this area.

3. A PARTICIPATORY APPROACH TO DEVELOPING CITIZEN SENSING

During 2013 to 2015 the Citizen Sense research project held a series of discussions and monitoring events with participants in northeast Pennsylvania, and from this process together developed the Citizen Sense Kit for monitoring air quality. The Citizen Sense Kit was used by a wide range of residents living near infrastructure, and also taken up by a local group, Breathe Easy Susquehanna County (BESC), interested in maintaining constructive dialogue with industry about changes in the environment particularly in relation to air quality. After a period of developing the Citizen Sense Kit in dialogue with participants, the Citizen Sense research project then deployed the kits in October 2014 with a training workshop and walk to field-test the technologies. The Citizen Sense research team then undertook visits to participants' homes to help set up the technologies, and participants developed a number of situations and experiments to monitor areas of particular concern to them.

The Citizen Sense kit distributed during these events has been has been developed in response to the concerns of community members, who provided information via Citizen Sense "logbooks" that asked for input on what the key concerns were for natural gas infrastructure in relation to air pollution. The Citizen Sense Kit itself, which was distributed to around 30 participants, contains a passive sampling badge for monitoring BTEX emissions (or benzene, toluene, ethylbenzene and xylene, which are substances associated with gas production and that are also hazardous to human health); along with a "Speck" device for monitoring particulate matter (PM) 2.5 (and these devices were loaned to us from the Create Lab at Carnegie Mellon). The kit also includes a custom-made prototype device, the Frackbox, which was installed at three compressor station sites. The Frackbox runs off a RaspberryPi and includes sensors for nitrogen oxide (NO), nitrogen dioxide (NO2), ozone (O3), and benzene, toluene, ethylbenzene, and xylenes (BTEX), as well as humidity and temperature. Participants were able to upload the data they gather to the Citizen Sense Kit platform, and to refer to a Citizen Sense logbook with instructions for use of the various part of the kit.

The kit attempts to provide ways for participants to document pollution events and experiences, and to observe patterns and relations that emerge from collected data. Citizen sensing often takes place in locations that are inaccessible to regulators and scientists, and the low-cost nature of the devices means the technologies can be distributed in a way that it is economically and practically impossible to install larger high-end devices.

In the pollution-sensing area of the Citizen Sense research project, this has been important for a number of reasons. Due to the sensitive context where environmental monitoring of fracking infrastructure is taking place, many of our participants needed to take part anonymously, as fears of reprisals from neighbours and industry were very present. The small size of the kit meant that it could be installed and used in an inconspicuous way, leading to citizens installing sensors in their porches, gardens, homes, and farmlands for three to six months. As participating citizens also lived distributed across the locality, instead of two or three monitoring points across an entire state, over 20 individual data points were created, which gave rise to the possibility of identifying localised sources of emissions, when read together with state air quality data.

The Citizen Sense Kit for monitoring air quality did not just focus on the gathering of numeric data, however. Photographs, mobile phone videos, YouTube comments, FLIR camera footage, diaries, and multiple other forms of documentation that on one hand might have seemed like a disparate set of resources, all contributed to the making of a “just good enough” collective data set for the region. Some participants began to notice patterns in their own data, as they uploaded the data to the CSK platform. Using the data together with wind speed and direction data from Weather Underground, participants were able to rule out spikes in their data that were most likely caused by regional sources and instead focus their energies on pollution events over more than 6 hours at time when the wind speed was lower suggesting a more local source. The participants who knew each other also formed sensing constellations and compared their data to each other.

Participants were further able to use their local expertise about fracking processes and infrastructure, in this case compressor stations, to answer the questions of the regulatory bodies, who had little day-today experience of living so close to natural gas extraction infrastructure. Another participant set up two monitors at a site opposite to a location that was scheduled to be fracked. Due to the temporal and unpredictable nature of much fracking, it has been difficult for regulators to monitor a well pad from start to completion. Companies may have a permit to drill a well for five years, and often may start fracking without warning. Therefore, one participant who passed a potential well pad site daily was able to establish a period of monitoring data before the fracking took place, and during the fracking. This combined with her daily YouTube videos of the drilling and fracking taking place, meant that there is now a unique set of evidence that can be read together.

Although in the view of the regulators the data generated by the Pollution Sensing project was not comparable to AQI air quality data, it was however “just good enough” for the participants to read together with state-collected air quality data and locally collected wind data from the platform Weather Underground. The distribution of devices also contributed to recognising a regional source of PM 2.5 in the area, which was good enough to form a pattern that could be excluded when looking at the local sources. One device on its own would probably not have been “just good enough,” but the distribution of devices, cared for by participants on a day-to-day basis over six months, made the data useful for entering into discussion with regulators, since in some cases even regulators and industry are unsure what is being emitted from these sites of concern.

Data that emerged through these techniques then became a useful negotiation tool, to arrange a number of conference calls with local bodies such as the Center for Disease Control and Prevention (CDC), the Agency for Toxic Substances and Disease Registry (ATSDR), the Pennsylvania Department of Health (DOH), and nonprofit environmental organisations as well as local political representatives. Although responses to the citizen-collected data ran the spectrum from outright dismissal to interest, there was just enough evidence to lead to one environmental agency requesting that local monitoring be undertaken, something which BESC participants had been campaigning for since the inception of their organisation.

4. DISCUSSION AND CONCLUSIONS

Although some citizen-sensing projects have worked closely with regulators and scientific disciplines, many others have departed from these practices, and instead use devices in unconventional ways, creating infrastructures that might be very different both spatially and temporally from those of the regulators. As citizen science and citizen sensing become more stabilized there is a call for practices to become more standardized and adopt generalized approaches to enhance the legitimacy of citizen-monitoring efforts. Indeed both the North American Citizen Science Association (CSA) and the European Citizen Science Association (ECSA) cite the need to establish standard protocols as central to the aims of the organizations. Gatherings such as the ECSA assembly are specifically coming together to form working groups to address this problem.

Much of the ongoing debate by practitioners and organizations that we have observed in citizen science meetings focuses on the importance of developing practices that can be directly comparable to existing regulatory practices. To some extent, there is a gap between the current sensing infrastructure and this vision. This has led to a drive towards designing devices that create data in similar formats, and to the calibrating of devices in reference to regulatory monitoring equipment. In some cases we have observed regulators recommending that citizens should only monitor in scenarios that are pre-approved by official bodies. Yet in this context, the inevitable question arises as to what new possibilities might be missed by attending only to the ways in which citizen sensing practices might replicate monitoring practices focused on regulatory compliance.

Citizen-gathered data using next-generation environmental sensing might have multiple uses, and the trajectories of citizen-sensing initiatives in making connections from environmental data to action do not need to exclude data that does not conform to regulatory practices, or which might have a more speculative starting point. This approach implies that any production of data by citizens that does not aim toward regulatory targets and processes cannot be useful. We recognize that for some contexts these new arrangements of infrastructures have proved challenging to both regulators and scientists, whose disciplines and professions have established practices of measuring, monitoring and accounting for environments. This in turn has often resulted in creating points of tension and disagreement between regulators and citizens.

But confining citizen sensing to conform only to regulatory standards would be to exclude the other creative and political possibilities of what we are calling "just good enough" data. And to align data practices exclusively with regulatory modes of monitoring might even exclude citizens from any participation in citizen sensing completely. For instance, in the context of air quality monitoring for PM 2.5, in order to be exactly comparable to the state DEP and US EPA air quality data, citizen-data would need to be collected by official, government trained personnel on approved equipment. Further to this, in order to be exactly comparable to regulatory data, the monitoring would have to take place at the very same locations, height and position that federal monitoring is already situated. In the context of the AQI PM 2.5 monitoring, this means monitoring would also have to be done over three years, and one could argue that as the data analysis process is also one of many decisions, data would have to be analysed (including averaged and smoothed) using the same software and algorithms as state and federal processes. In this scenario, citizen sensing as a practice would become completely redundant as the process would have to exactly replicate the monitoring already being performed by governmental agencies and experts, rather than opening up opportunities for monitoring to be undertaken by a wider range of participants, in varied locations, over different timescales, and in response to distinct events.

Instead, we suggest that "just good enough" data, while not ignoring the important issues of accurate instrumentation, calibration, and measurement, along with robust monitoring practices, might also allow more expansive uses of citizen-sensing technologies and data, while still opening up a dialogue on environmental change between citizens and regulators. With this proposal, we are not regressing to early conceptions of public science, where the collection of data is a cursory one to raise public awareness, but on the contrary we suggest that "just good enough data" is a practice that creates a shared space for discussion that can communicate community awareness of pollution events to regulators. Citizen-produced data sets often declared to be inaccurate due to the devices used, illegitimate due to the protocols followed, and unscientific due to the perceived community bias (such as citizens monitoring to create deliberate evidence for pollution events). However, we have shown that citizen sensing is also an entry point for testing the claims about the ease of participation that next-generation environmental sensors are meant to offer, as well as of developing expanded aspects of participation, monitoring, and environmental politics, which might allow communities to engage more readily with environmental problems.

5. ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313347, "Citizen Sensing and

Environmental Practice: Assessing Participatory Engagements with Environments through Sensor Technologies."

6. REFERENCES

Breathe Easy Susquehanna County. https://www.facebook.com/BreatheEasySusq.

Citizen Sense. http://www.citizensense.net.

MarcellusGas.Org. https://www.marcellusgas.org.

State Impact. "The Marcellus Shale, Explained."

U.S. Environmental Protection Agency. "Draft Roadmap for Next Generation Air Monitoring." March 8, 2013. http://www.eunetair.it/cost/newsroom/03-US-EPA_Roadmap_NGAM-March2013.pdf.

Engaging Citizens in Environmental Monitoring via Gaming

Ian McCallum[1], Linda See[1], Tobias Sturn[1], Carl Salk[1], Christoph Perger[1], Martina Duerauer[1], Mathias Karner[1], Inian Moorthy[1], Dahlia Domian[1], Dmitry Schepaschenko[1], Steffen Fritz[1]

[1]International Institute for Applied Systems Analysis

mccallum@iiasa.ac.at

Abstract

Citizen science is quickly becoming one of the most effective tools for the rapid and low-cost collection of environmental information, filling a long recognized gap in in-situ data. Incentivizing citizens to participate however remains a challenge, with gaming being widely recognized as an effective solution to overcome the participation barrier. Building upon well-known gaming mechanics, games provide the user with a competitive and fun environment. This paper presents three different applications that employ game mechanics and are generating useful information for environmental science. Furthermore, it describes the lessons learnt from this process to guide future efforts.

Keywords: Crowdsourcing, citizen science, gaming, land cover, land use

1. INTRODUCTION

Citizen science is increasingly being recognized as an important new, component of environmental monitoring. There are many different examples of successful, ongoing citizen science projects (http://scistarter.com/) but many of these projects are focused on biodiversity, species identification and nature conservation. There are fewer examples of citizen science projects that involve citizens in the collection of data for the calibration and validation of Earth Observation (EO) products, yet the potential clearly exists. One example of such a project is Geo-Wiki (Fritz et al., 2012), which is an online tool that is focused on gathering calibration and validation data of land cover from high resolution satellite imagery. A number of crowdsourcing campaigns have been organized in the past, where crowdsourced land cover data have been used to produce hybrid land cover products and to validate existing products (see e.g. Fritz et al. (2013, 2015); See et al. (2015)). Although these campaigns were successful both in terms of the quality of data collected and the involvement of people in science, we wanted to investigate methods for attracting larger numbers of participants and thereby develop a much larger training and validation database. One way to achieve this goal is via gaming and the use of mobile devices, since games are the most frequently used application on smartphones (dotMobi, 2013).

The use of serious games (or games with a scientific purpose) is not new. There are many good examples, some of which have shown impressive scientific results. For example, FoldIt is a game in which players fold proteins and has led to discovering new protein strucutres (Khatib et al., 2011). A game to help map the neurons in the brain, called EyeWire (Kim et al., 2014), is being played by more than 160,000 people. In this paper we present three serious games in the context of land cover / land use monitoring that have been developed as part of the current set of available Geo-Wiki tools along with lessons learned.

2. GEO-WIKI SERIOUS GAMES

2.1. Cropland Capture

Cropland Capture was launched in mid-November 2013 and ran until the beginning of May 2014 as a multi-platform game running in a browser, on a smartphone or a tablet for both the Apple and Android operating systems. As part of the game the players were presented with a red rectangle placed on top of satellite imagery or photographs. Players were then asked to determine if there was any evidence of cropland in the image. The mobile device interface was designed so that players swipe the images into three possible categories of Yes, No or Maybe. For each correct answer, the player received a single point while one point was deducted for incorrect answers. Correctness was determined through majority agreement although there was an option to challenge the crowd. The leaderboards were reset each week and the top three players in terms of number of classifications were added to a prize draw that took place at the end of the six month period.

At the end of the game, more than 4.5 million images and photographs had been classified. Of these, there were around 170,000 unique images. This means that each image has a frequency distribution, which has allowed us to examine the performance of the crowd. Overall users disagreed with the crowd less than 10% of the time, with low bias toward identification of cropland or non-cropland. This implies that identifying cropland from the images and photographs is a task that the crowd can easily do with high accuracy. There was also no significant difference in performance based on background or experience, implying that a background in remote sensing is not required for satisfactory performance in this task. Other patterns found in the data are described in Salk et al. (2015). More information about the game can be found at: http://www.geo-wiki.org/oldgames/croplandcapture/

2.2. Foto Quest Austria

IIASA launched Foto Quest Austria in July, 2015, running for three months. The app has the EU LUCAS (Land Use and Cover Area frame Survey) protocol built-in and asks players to locate points on the ground, classify the land cover based on the LUCAS categories and then take pictures in 4 directions and on the ground. The game-like app, which is aimed at a German-speaking audience of all ages, allows participants to take geo-located photographs of landscapes for science.

The project aims to gather information about land use change in Austria that is important for research on climate change and flood risk. It also aims to spark a sense of adventure and exploration, encouraging participants to get outside and enjoy nature. Participants can also compete for points and prizes. In particular, we are interested in gathering data on urban expansion and the preservation of wetlands, which store large amounts of carbon dioxide and are therefore important for limiting climate change.

The game collected a total of over 12,000 photos, with more than 2000 locations visited or 12% of total available locations. Approximately 200 persons downloaded the app and collected a minimum of one observation. We are currently working with the TD1202 COST network Mapping and the Citizen Science, to use the app in other European countries. More information can be found at: http://fotoquest.at/

2.3. Picture Pile

The Picture Pile game is the successor to Cropland Capture. The game has been made more generic, i.e. other land cover types have been added and it will also collect information on land use as well as change over time. The idea is that players classify piles of pictures, where a pile represents one task or theme. By having different tasks, more variety will be embedded in the game to help retain players for longer. The game mechanics are similar to Cropland Capture, i.e. the player swipes images to the Yes, No or Maybe categories to classify them.

Some other planned tasks include the following: ‘Do you see evidence of human impact in this picture?’ and ‘Are the trees oil palms?’ For change detection, players will be presented with a pair of images and asked to determine whether they can spot any evidence of deforestation. Other tasks will be embedded as the game progresses. Each pile has an associated leaderboard and a chat channel, which makes it very easy for players and the organizers to communicate and will foster the formation of a community.

The scoring mechanism that will be employed in Picture Pile is one of the main changes from the Cropland Capture game. Instead of a majority agreement approach, more use of controls or reference data will be implemented within the game. The reference data will be images that experts have interpreted and are therefore treated as correct. Reference data will be provided to the players throughout the game and incorrect answers will be heavily penalized. This mechanism will also be used to provide feedback to the players so that they learn over time as the game runs. The game will be launched in November 2015 and more information can be found at: http://geo-wiki.org/games/picturepile

3. DISCUSSION AND CONCLUSIONS

This paper has briefly described three game applications that are generating considerable amounts of in-situ data for terrestrial environmental monitoring. Not only is the volume of data impressive, but the cost at which it is obtained is far less than collecting authoritative data. In addition, the nature of crowdsourcing or citizen science means that information can be obtained over time allowing for new forms of change detection.

One of the most challenging aspects of citizen science includes overcoming the barriers to participation. From the gaming applications described above, a variety of lessons have been learnt, which could apply generally to game applications across numerous themes. These include the need to develop a good, professional looking app, which provides personal satisfaction to the user and has a fun feel to it. However, a good looking app alone is not sufficient; it must be a part of a holistic approach that contains stimulating, scientific ideas, allows for good communication with the community and provides appropriate incentives for participation.

Testing the app is critical so a reliable testing strategy should be built into the app development. For researchers entering this field, the amount of testing needed should not be underestimated. Moreover, developing for Android can be tricky as there are many different Android versions and phones available. However, the app should be as bug free as possible or users will not engage further. Testing is critical not only for the first release, but for all subsequent updates as well, since new bugs can be introduced.

Campaigns require considerable thought when using volunteers to do the work of expert (authoritative) data collection. There must be a lot of tips and helpful hints on how to do what is needed without overburdening the users (and thereby losing their interest). The experience with the collection of data following a LUCAS protocol demonstrated the need for many easy-to-follow tips and hints.

The timing of a campaign launch should be carefully chosen to match the target audience, e.g. if targeting schools, then campaigns should be integrated within the school year. Media launch events, i.e. press conferences, interviews and video for high-level media distribution, are important for initial and ongoing recruitment and outreach. Another key to success is to be well connected to grassroots organizations (e.g. the Alpenverein in the case of Foto Quest Austria). Outreach via social media is also an effective channel. Including a budget line for facebook-targeted campaigns that generate direct website clicks, mobile installs, etc. can yield additional participation.

In conclusion, engaging citizens in environmental monitoring via gaming as demonstrated by the three examples given here and numerous other activities underway has great potential to change the way we collect in-situ data. Nevertheless, numerous pitfalls exist with lessons learnt from previous game development being applicable across many domains. In spite of the challenges outlined here, gaming applications offer the possibility to generate large amounts of timely, cost-efficient high quality information that was previously unavailable.

4. REFERENCES

dotMobi (2013). *Global mobile statistics 2013 Section E: Mobile apps, app stores, pricing and failure rates*, at http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats/e#popularappcatagories, [accessed 7 March 2014](2013). Global mobile statistics 2013 Section E: Mobile apps, app stores, pricing and failure rates.

Fritz, S., I. McCallum, C. Schill, C. Perger, L. See, D. Schepaschenko, M. van der Velde, F. Kraxner and M. Obersteiner (2012). Geo-Wiki: An online platform for improving global land cover, *Environmental Modelling & Software*, 31: 110–123.

Fritz, S., L. See, I. McCallum, L. You, A. Bun, E. Moltchanova, M. Duerauer, F. Albrecht, C. Schill, C. Perger, P. Havlik, A. Mosnier, P. Thornton, U. Wood-Sichra, M. Herrero, I. Becker-Reshef, C. Justice, M. Hansen, P. Gong, S. Abdel Aziz, A. Cipriani, R. Cumani, G. Cecchi, G. Conchedda, S. Ferreira, A. Gomez, M. Haffani, F. Kayitakire, J. Malanding, R. Mueller, T. Newby, A. Nonguierma, A. Olusegun, S. Ortner, D.R. Rajak, J. Rocha, D. Schepaschenko, M. Schepaschenko, A. Terekhov, A. Tiangwa, C. Vancutsem, E. Vintrou, W. Wenbin, M. van der Velde, A. Dunwoody, F. Kraxner and M. Obersteiner (2015). Mapping global cropland and field size, *Global Change Biology*, 21(5): 1980–1992.

Fritz, S., L. See, M. van der Velde, R.A. Nalepa, C. Perger, C. Schill, I. McCallum, D. Schepaschenko, F. Kraxner, X. Cai, X. Zhang, S. Ortner, R. Hazarika, A. Cipriani, C. Di Bella, A.H. Rabia, A. Garcia, M. Vakolyuk, K. Singha, M.E. Beget, S. Erasmi, F. Albrecht, B. Shaw and M. Obersteiner (2013). Downgrading recent estimates of land available for biofuel production, *Environmental Science & Technology*, 47(3): 1688–1694.

Khatib, F., F. DiMaio, F.C. Group, F.V.C. Group, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska, I. Pichova, J. Thompson, Z. Popović, M. Jaskolski and D. Baker (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players, *Nature Structural & Molecular Biology*, 18(10): 1175–1177.

Kim, J.S., M.J. Greene, A. Zlateski, K. Lee, M. Richardson, S.C. Turaga, M. Purcaro, M. Balkam, A. Robinson, B.F. Behabadi, M. Campos, W. Denk, H.S. Seung and The EyeWirers (2014). Space-time wiring specificity supports direction selectivity in the retina, *Nature*, 509(7500): 331–336.

Salk, C.F., T. Sturn, L. See, S. Fritz and C. Perger (2015). Assessing quality of volunteer crowdsourcing contributions: lessons from the Cropland Capture game, *International Journal of Digital Earth*, : 1–17.

See, L., S. Fritz, C. Perger, C. Schill, I. McCallum, D. Schepaschenko, M. Duerauer, T. Sturn, M. Karner, F. Kraxner and M. Obersteiner (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*, 98: 324–335.