International Journal of Spatial Data Infrastructures Research, 2018, Vol.13, 78-77
Special Section: Citizen Data Science

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], Paul van Genuchten[6], Michael O'Grady[7]

[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]GeoCat, [7]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

1. INTRODUCTION

New and innovative environmental monitoring and information capabilities can enable citizens to effectively participate in environmental monitoring, based on

---

**This paper is a shorter version of a previous publication in this journal (DOI: 10.2902/1725-0463.2016.11.art3). It was produced for a dedicated workshop on the topic of Citizen Science.

DOI: 10.2902/1725-0463.2018.13.art8

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 the 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; and 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 who are 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.

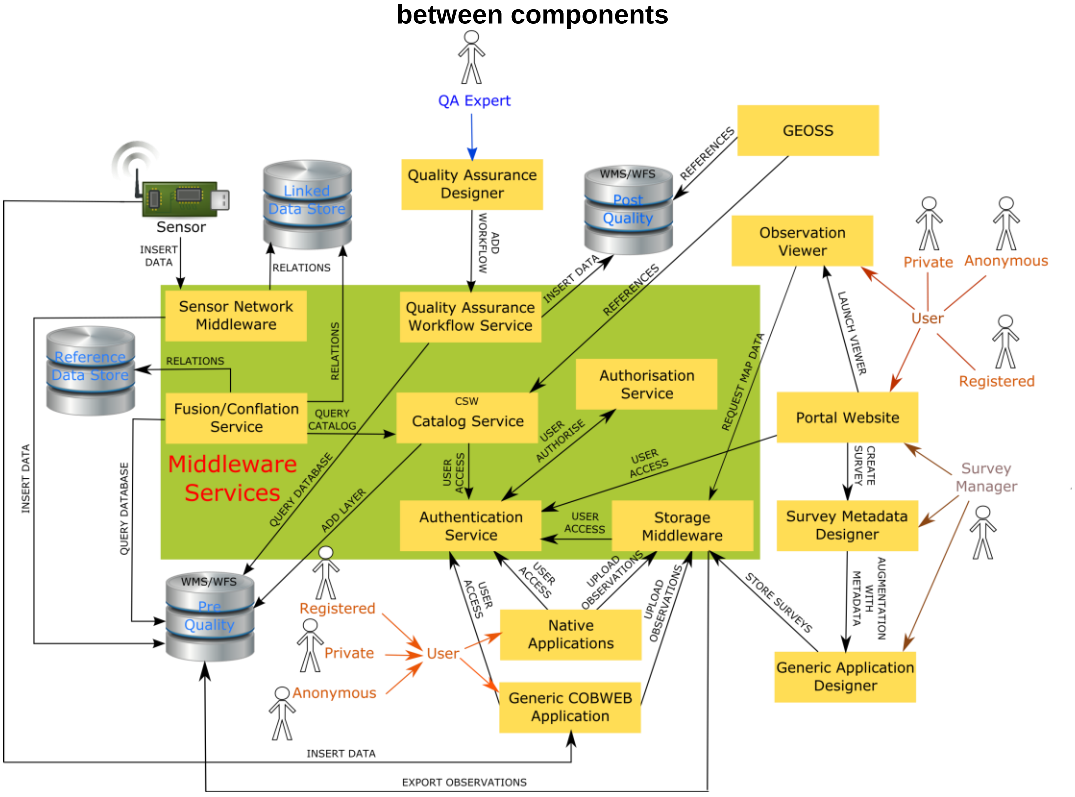
2. SYSTEM DESIGN THROUGH STAKEHOLDER ENGAGEMENT AND CO-DESIGN

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.

3. ARCHITECTURE

The COBWEB system architecture (Figure 1) has developed through a combination of new software builds which resulted from periods of rapid prototyping and which built upon open source software; however, the system architecture continues 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. It also implements access control and privacy as well as open standards. Together these offer 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



It is important to note that the various components are at different levels of technical maturity ranging from Technology Readiness Level (TRL) 4 (validated in lab) to TRL 9 (actual system proven in operational environment). The TRL scale has been developed to enable assessment and comparison of technologies with respect to maturity and is used as a tool for decision-making on research and development investments at the EU level (EARTO, 2014).

3.1 Portal Website

An implementation of the architecture illustrated in Figure 1 was developed in association with the Dyfi Biosphere Reserve in mid-Wales. A portal website accessible via the main Dyfi Biosphere Reserve website was engineered as the main point of entry to this instance of the COBWEB framework.

At the centre of the portal is the latest version of GeoNetwork, improved to facilitate citizen science. GeoNetwork is a catalogue application which offers registration capabilites for resources and exposes those resources via the Open Geospatial Consortium's (OGC) Catalogue Services for the Web standard (Nebert et al., 2007), OAI-PMH, ATOM/OpenSearch and html. Main input and storage schemas are ISO19139 and ISO19115-1, but alternative output schemas are available, such as Dublin Core, SensorML, DCAT, PPSR_CORE and schema.org. This version of GeoNetwork supports the concept of registration of 'surveys,' also known in the community as 'citizen science projects' in (PPSR_CORE, 2015).

Users can login to the portal and request to join selected citizen science surveys. 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 using Web Map Services (WMS), Web Feature Services (WFS) or Sensor Observation Services (SOS).

3.2 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 which enables survey managers to build custom data collection forms using the generic application designer component of the portal website. Forms designed at the portal can then be synchronised with the generic COBWEB application on individual users' mobile devices (Butchart, 2013). The generic application designer supports a wide variety of form elements to cater to a broad range of user requirements.

3.3 Apps

The generic App (an application, downloaded by a user to a mobile device) approach adopted, provides the citizen with the abilty to install the COBWEB App 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 to 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 to show 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).

3.4 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.

3.5 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, their 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 (the QA Workflow Service in Fig 1), This allows great flexibility in what quality processes are applied to the citizen science data captured. This generic capability of authoring and customising QA is necessary because exactly what quality control processes are applied in a particular situation is often highly use case dependent.

Besides allowing reuse of other web services within the QA workflow, the solution researched extends a pre-existing typology of quality assessment types (Goodchild & Li, 2012) to seven categories (or pillars) covering a range of specific quality controls generating quality metadata elements: 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 (WPS) 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) has been customised to work with OGC services (Meek et al., 2015) and has been integrated as a component of the portal website. The Survey Manager has the authority to create quality assurance workflows.

Data fusion and data conflation in COBWEB is used either after QA for final data use or during QA for some validation of observations with external resources available on the Web (Wiemann et al., 2015; Leibovici et al., 2015). Research has focused on making the functionality accessible through the QA Workflow Service portal component. 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 & Bernard, 2015). Access is also via WPS interfaces. Provenance is registered as within a linked data store.

3.6 Sensor Networks

Though data collection from mobile devices is fundamental to COBWEB, a variety of different sensor platforms monitoring multiple environmental parameters within the Dyfi Biosphere Reserve testbed, have been deployed. These sensor networks have been specified in accordance with 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 QA 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, data sources external to COBWEB can be integrated into an arbitrary survey when data are made available in a public and standards-compliant way. It is envisaged that, going forward, public authorities will increasingly make selected data sources available in this way.

3.7 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 detailed information on these species available over the web. Conversely, these are often the most valuable data for consideration in environmental monitoring and is 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 as well as 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.

3.8 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 advancing 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.

4. SUMMARY

COBWEB is conducting research into the feasibility of creating a common framework for mobile device apps for use in citizen science for environmental monitoring, using the UNESCO World Network of Biosphere Reserves as a testbed.

It has been 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 research is needed to address whether it is possible to create a useable framework which is sufficiently flexible to allow a broad range 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 efforts should continue and adherence should be required to help realise investment in spatial data infrastructure type initiatives such as GEOSS.

The components of the developed solution are at differing TRLs and continue to evolve as the project progresses. At project completion, where applicable, the source code will become available to the wider community for further development.

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 Koutroumpas, M. (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]

EARTO (2014) The TRL Scale as a Research & Innovation Policy Tool, EARTO Recommendations, at http://www.earto.eu/fileadmin/content/03_Publications/The_TRL_Scale_as_a_R_I_Policy_Tool_-_EARTO_Recommendations_-_Final.pdf [accessed 7 Jane 2016]

Goodchild, M.F., and 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., and 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, M., Broday, D., and Bartonova, A. (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, D.G. (2015) Addressing the quality assurance challenge for location-based crowd-sourced data through workflow composition of OGC web services. Computers & Geosciences (submitted)

Nebert, D., Whiteside, A., and Vretanos, P (Eds). (2007) OpenGIS Catalogue Services Specification version 2.0.2, Open Geospatial Consortium

PPSR_CORE (2015) Citizen Science Association http://citizenscienceassociation.org/2015/06/11/efforts-underway-to-share-project-metadata/ [Accessed 8 Jan 2016]

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

Wiemann, S., Meek, S.M., Chapman, C., Leibovici, D.G., Jackson, M., 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].