Joan Masó (CREAF), Ester Prat (CREAF), Andy Cobley, (University of Dundee)

EXPERIENCES USING SOS FOR CITIZEN SCIENCE INTEROPERABILITY

September 2–4, 2019 Münster, Germany

Geospatial Sensing Conference 2019
Imagine a world where 503 environmental science projects collect data in 503 independent silos.
Terrified?
...and yet
this is the current situation
in Citizen Science
De-silo (verb)

de-silo, de-siloing, de-siloed verb, transitive

- To integrate data from disparate sources maintained by separate departments.
  To *de-silo* means to get rid of silos

From: [https://www.pcmag.com/encyclopedia/term/70006/de-silo](https://www.pcmag.com/encyclopedia/term/70006/de-silo)
You can do many things to de-silo:

- Define common data models and vocabularies
- Set up bridges between organizations
- Upload the data in a single repository
- Draft comprehensive data management strategies

...or

- Setup web services to make data accessible and well documented
In the first OGC Citizen Science Interoperability Experiment (CitSciIE)

- The use of OGC standards (e.g. Sensor Web Enablement (SWE)) to support **data sharing and documentation** among Citizen Science projects, and combine it with other sources, (e.g. authoritative data)
Following SWE4CS

- A publicly available discussion paper in the OGC mainly produced by the FP7 COBWEB project
Thanks to the initiators and supporters
Focus on Data sharing using OGC standards such as O&M and SOS

- We had four SOS services

- We had three SOS clients
...and a part from good meetings...

Stuttgart
Contribution to the #CITSciE: @opengeo spatial definitions server to store the concepts. Extended to the ones used by the projects.

Kyoto
Joan Masó Pau
@joanma747
She mentioned the opengeospatial as partner and the #CItSciE interoperability experiment we are doing in !WeObserveEU

Venice
Joan Masó Pau
@joanma747
We detected and issue on the position of a moving sensor (common in cisci) versus the position of the observation in SOS C&M. How to link observations with sensors and procedures? Discusses in the #CitSciE @FivoTwoN @opengeo spatial @COWM2018 @CREAF_ecologia @WeObserveEU

Vienna
Uta Wehn
@UtaWehn
@joanma747 presents the @GroundTruth20 data quality tool during @WeObserveEU # interoperability #CoP at #EGU19, applied to @MeetVoeMechelen & @hack_data. @EU_H2020
... we did some TIEs

(Technology Integration Experiments)
Using some data
We are wrapping up the activity

• The external twiki:

• A GitHub repo to do the Engineering Report
  - https://github.com/opengeospatial/CitSciIE
  - Will be finalized soon
Conclusion

Sensor Observing Service can be used to make Citizen Science data interoperable.

Demonstrated!
Architecture one
In action in a JavaScript client

http://www.ogc3.uab.cat/gt20/?config=hackair.json
Architecture two

Interoperability

Common server

SOS requests

SOS server

OGC

CitSci

HydroLogic

Dada

API 3
Architecture three

Interoperability

Helgoland

Helgoland

SOS requests

SOS requests

52n

52n

52n

Helgoland API 4 JSON
Helgolan SOS in action
How many APIs?

...but only one SOS
Spanish case: Phenological observations exposed as WFS

Experiment 1: What is behind?

• A OGC WFS GetFeature encoded in GeoJSON

• Maximum compatibility. There are more implementations of WFS than SOS

• We are not respecting the SWE4CS
Swedish case as SOS service

Experiment 2: What is behind?

- A OGC SOS GetObservation encoded in XML
- Carefully respecting the standard
- Only agile with a tens of points
Belgium case as SOS service

Experiment 3: What is behind?

- A OGC SOS getObservation encoded in JSON
- JSON is far more agile in web browsers and it can deal with thousands of points
- O&M in JSON is still not recognized as an official OGC standards even if there are some implementations
Some conclusions

• CitSci tend to produce several observations in different places
  • This creates many FeaturesOfInterest (many points) and many individual Observations generating very verbose files.
• Issue in real implementations: XML encoding is difficult to parse and JavaScript libraries are slow. A file containing 8000 observations and 10 variables can easily generate a 30Mb file that is huge for the DOM parsers.
  • This results in about 500000 individual calls to XML libraries (to navigate for each field extract the properties of the features) that can take from 20 seconds up to 5 minutes depending on implementations. The use of canvas to represent does 8000 points in the screen is far much faster.
Some conclusions

• Switching to WFS-GeoJSON and SOS-JSOM-”free-style” fixes the problem on the browser

• There is an need for making a JSON encoding for SOS and O&M official and popular

• An alternative is to use directly SensorThings API that already provides a JSON encoding also inspired on O&M
From OGC SOS to SensorThings API: #ScentEU #EGW2017 Athanasia Tsertou

<table>
<thead>
<tr>
<th>Standard / Data Model &amp; Implementation</th>
<th>OGC SOS (v2.0) - 52°North implementation</th>
<th>Senslog Data Model</th>
<th>SensorThings API – e.g. Fraunhofer implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>OGC SOS</td>
<td>Open-source, Apache 2.0</td>
<td>Open-source, BSD 3-clause Lightweight Rest API VGI module support Multimedia support</td>
<td>Open-source, GPL v3.0 IoT-like, recently standardised Compatible with OGC SISO M</td>
</tr>
<tr>
<td>REST API</td>
<td>Supported in GEOSS registry</td>
<td>REST API</td>
<td>Support Multimedia not supported per se; yet OGC Testbed-10 CCI VGI Engineering Report</td>
</tr>
<tr>
<td>Good documentation</td>
<td>Multiple Implementations</td>
<td>Not in GEOSS registry Limited implementations</td>
<td></td>
</tr>
<tr>
<td>Heavy, not scalable VGI not supported per se; yet OGC Testbed-10 CCI VGI Engineering Report</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Part of a continuous effort
Communities of Practice in WeObserve

CoP = Communities of Practice
Groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly

Domain (Relevance)
Community (Reflection)
Practice (Contribution)

Source: Yusuf et al. (2011)

KNOWLEDGE BASE

AWARENESS
COP 1: ENGAGING STRATEGIES & CO-DESIGN COs

ACCEPTABILITY
COP 2: IMPACT & VALUE OF COs FOR GOVERNANCE

SUSTAINABILITY
COP 3: INTEROPERABILITY & STANDARDS FOR COs

COP4: SUSTAINABLE DEVELOPMENT GOALS & CITIZEN OBSERVATORIES
Next Interoperability Experiment

- Test interoperability of CS project using **SensorThings API**
  - complementary with previous tests done with SOS

- Advancements on architectures for a federation and single entry point for CS projects
  - Serving to connect to big systems such as GEOSS, GEOBON etc.

- Developing vocabularies for Citizen Science and minimum common denominator data models

- Open to other ideas
  - We are defining this now.

- Starting in November in Barcelona (and Toulouse)
Open Data Challenge

Andy Cobley

University of Dundee
Open Data Challenge

**Brief Overview**
- Build on interoperability Experiment
- Aimed at SME and Start ups
- **Use data** from Citizen Observatories
- Combining other open sources

**How can you help**
- Nominate judges (August 2019)
- Nominate datasets (August 2019)
- Spread the word from October 2019

**Time Line**
- **January 2020**: Access to data sets
- **May 2020**: SME Demos
- June 2020: Award

**Contact**
- aecobleyn@Dundee.ac.uk
- s.m.coulson@dundee.ac.uk
THANK YOU!

Any Questions?

Joan Masó
CREAF
joan.maso@uab.cat