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

EXPERIENCES USING SOS FOR CITIZEN SCIENCE INTEROPERABILITY

Imagine a world where

503

environmental science projects

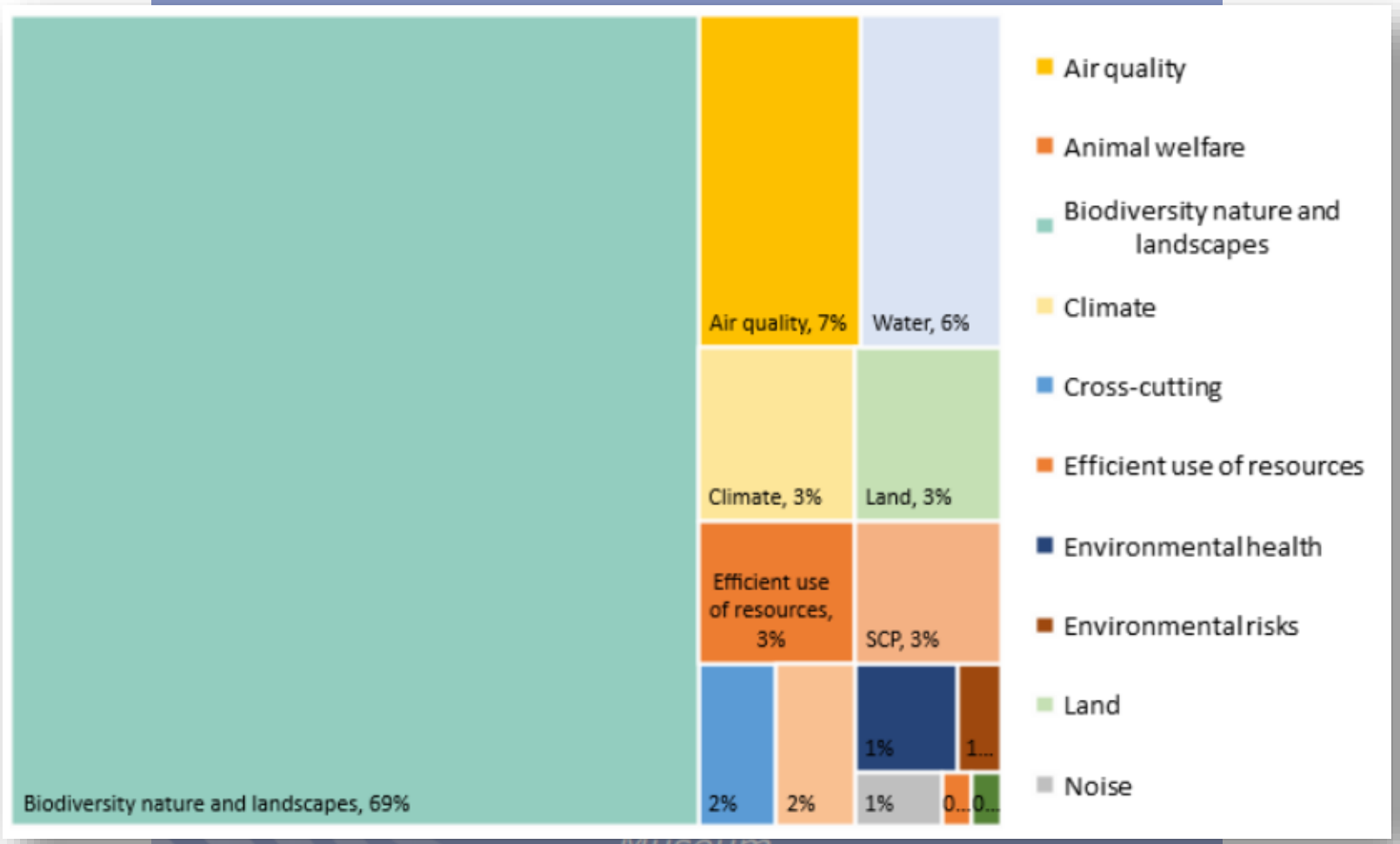
collect data in

503

independent silos


Terrified?

...and yet
this is the current situation
in Citizen Science



http://data.europa.eu/euodp/data/dataset/jrc-citsci-10004

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

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 document

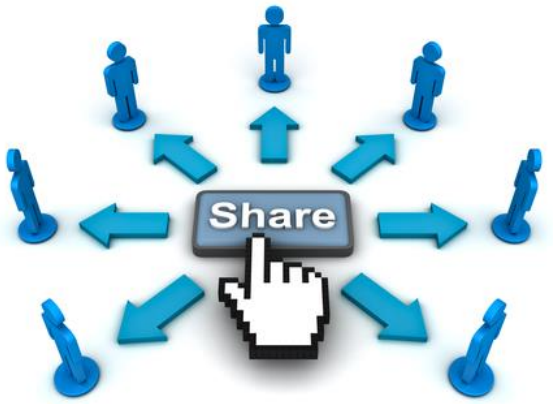
In the first OGC Citizen Science Interoperability Experiment (CitSciIE)

- 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)

CitSciIE

Following SWE4CS

- A publicly available discussion paper in the OGC mainly produced by the FP7 COBWEB project



Thanks to the initiators and supporters

Initiators

IIASA

JRC EUROPEAN COMMISSION

esa

Wilson Center

UAB Universitat Autònoma de Barcelona

CREAF

Support

Interop-CoP

We observe

NEXTGEOSS

OGC

Focus on Data sharing using OGC standards such as O&M and SOS

- We had four SOS services

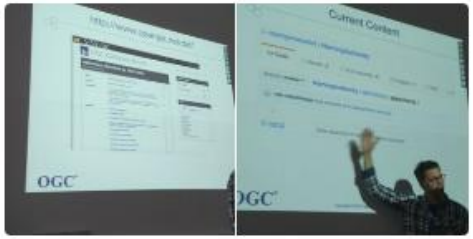
- We had three SOS clients

…and a part from good meetings…

Stuttgart

Joan Masó Pau
@joanma747

Contribution to the #CitSciIE: @opengeospatial definitions server to store the concepts. Extended to the ones used by the projects.



Sep 14, 2018

Joan Masó Pau
@joanma747

Contribution to the #CitSciIE: @LandSense engagement platform and single sign on.

Sep 14, 2018

Kyoto

Joan Masó Pau
@joanma747

nne is presenting the architecture of the data cloud at will support the Earth challenge 2020. GEOWeek18 . She mentioned the opengeospatial as partner and the #CitSciIE teroperability experiment we are doing in WeObserveEU

Oct 29, 2018

Venice

Joan Masó Pau
@joanma747

We detected and issue on the position of a moving sensor (common in citsci) versus the position of the observation in SOS O&M. How to link observations with sensors and procedures? Discussed in the #CitSciIE @FiveTwoN @opengeospatial @COWM2018 @CREAF_ecologia @WeObserveEU

Nov 28, 2018

Vienna

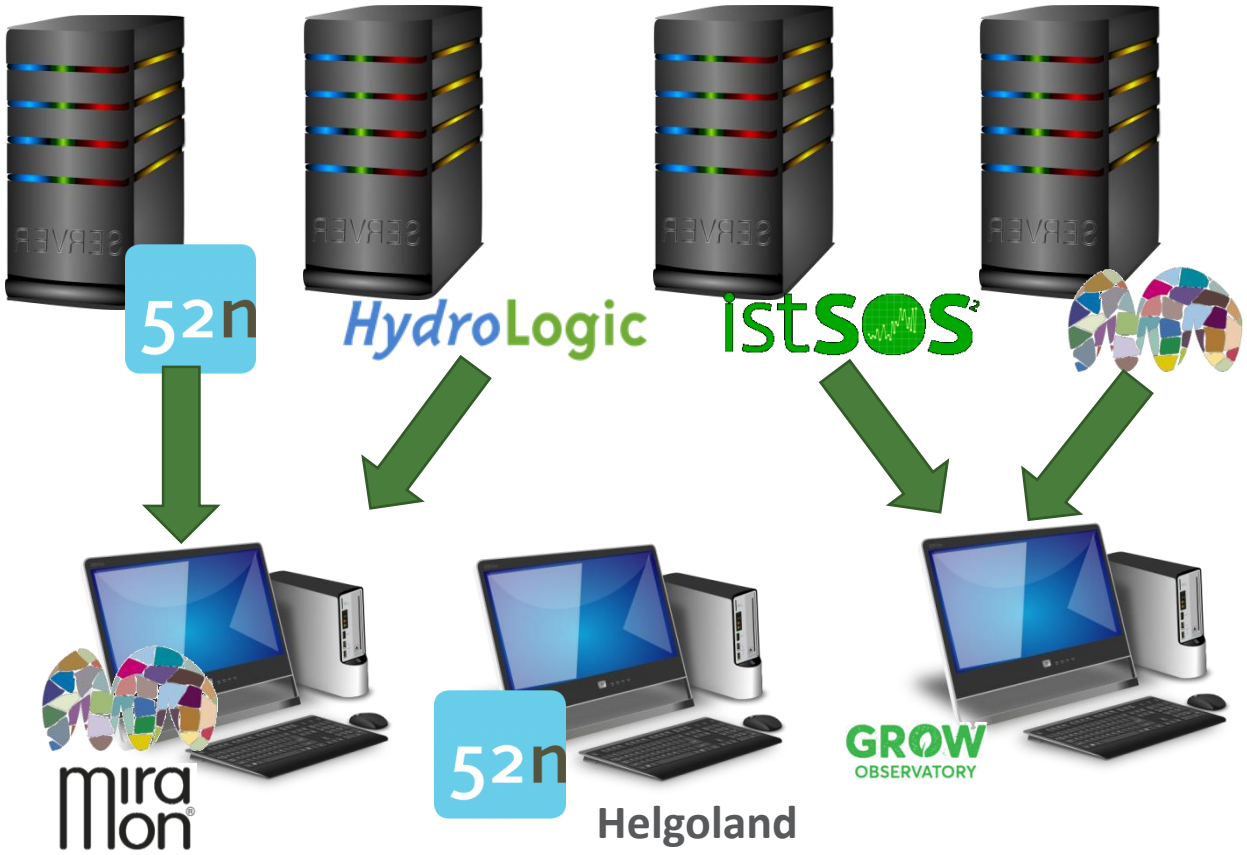
Uta Wehn
@UtaWehn

@joanma747 presents the @GroundTruth20 data quality tool during @WeObserveEU #interoperability #CoP at #EGU19, applied to @MeetMeeMechelen & @hack_air data. @EU_H2020

Apr 11, 2019

... we did some TIEs

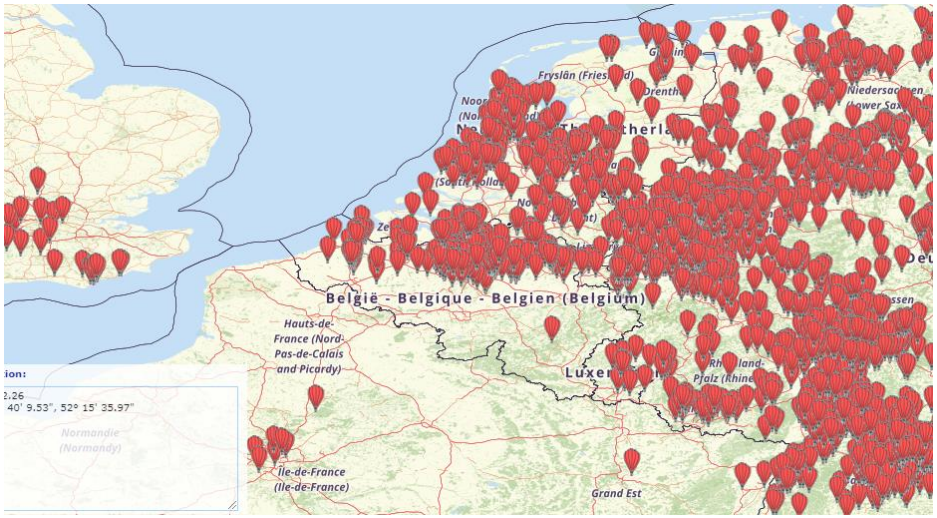
(Technology Integration Experiments)

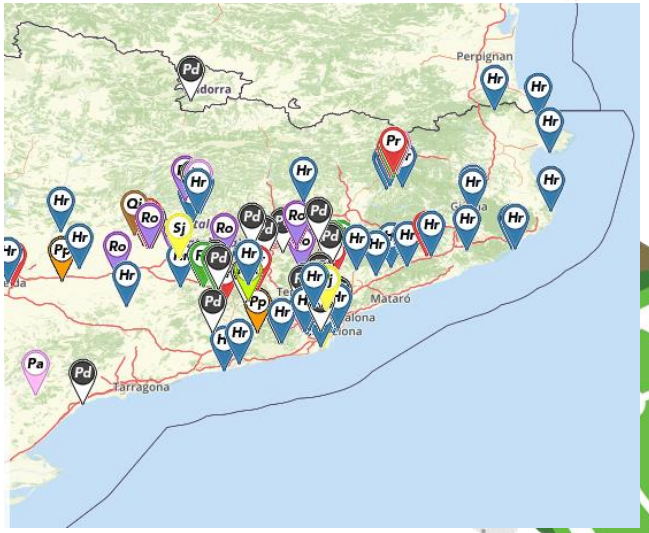


Using some data

hackAIR

groundtruth2.0
Environmental knowledge discovery of human sensed data
RitmeNatura.cat

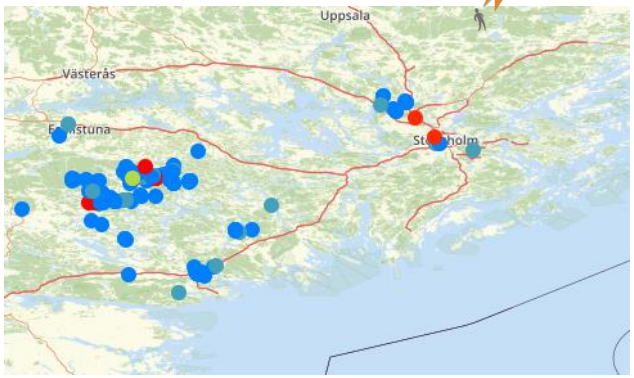


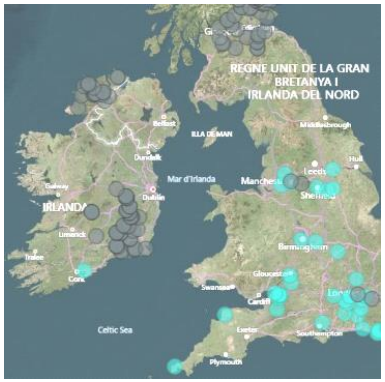


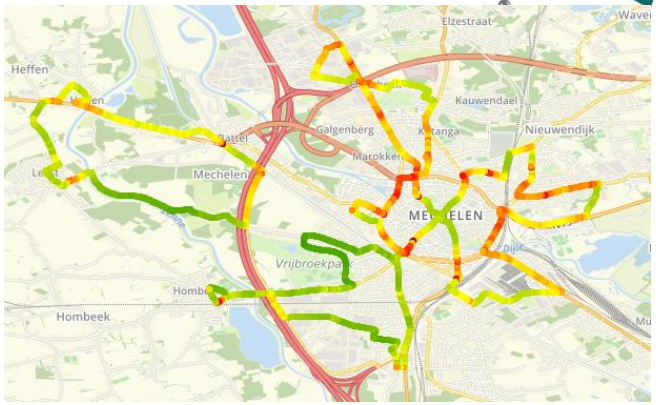
groundtruth2.0
Environmental knowledge discovery of human sensed data
VattenFokus

GROW OBSERVATORY

groundtruth2.0
Environmental knowledge discovery of human sensed data
Meet Mee Mechelen







We are wraping up the activity

- The external twiki:
  - http://external.opengeospatial.org/twiki_public/CitSciIE/WebHome

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

CitSciIE

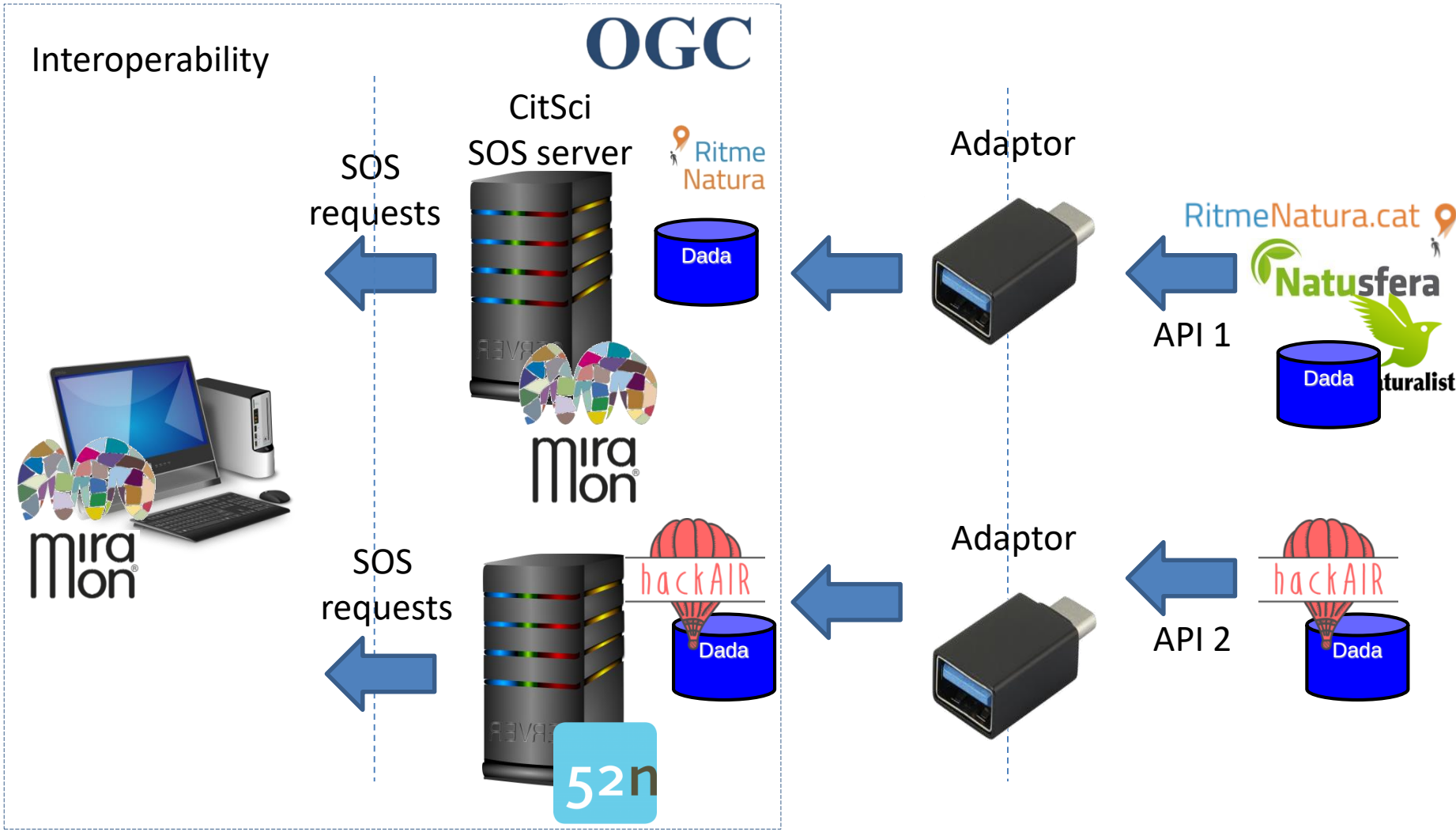
Conclusion

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

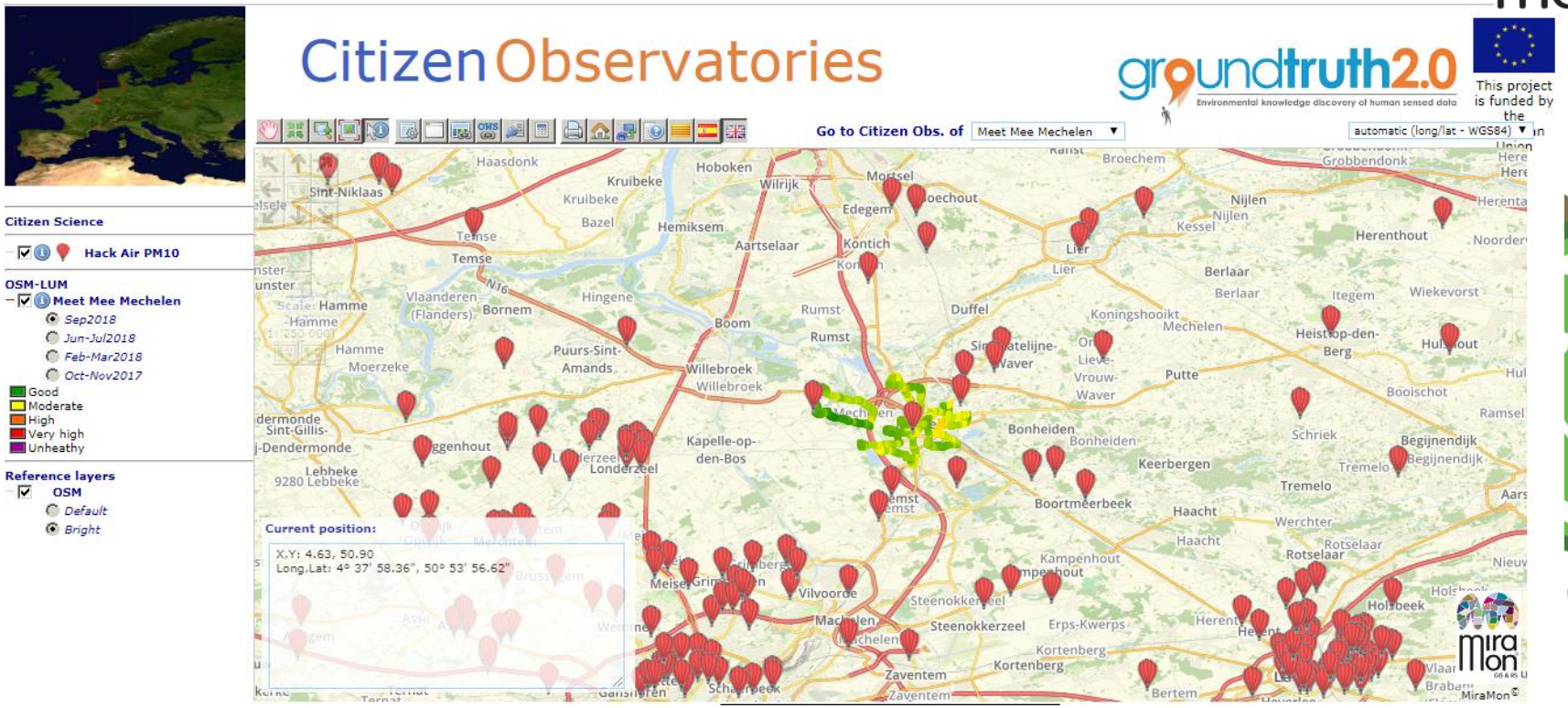
Demonstrated!

Architecture one

①



In action in a JavaScript client

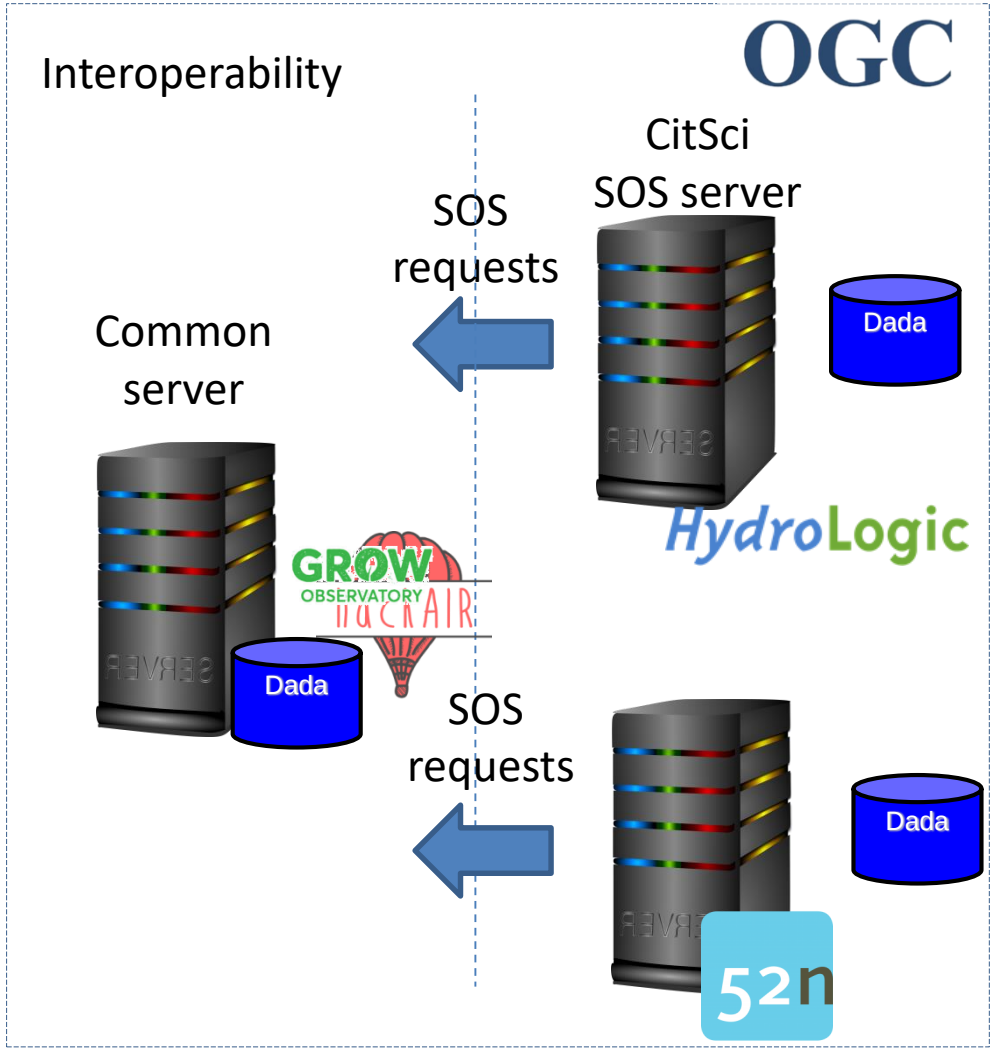


http://www.ogc3.uab.cat/gt20/?config=hackair.json

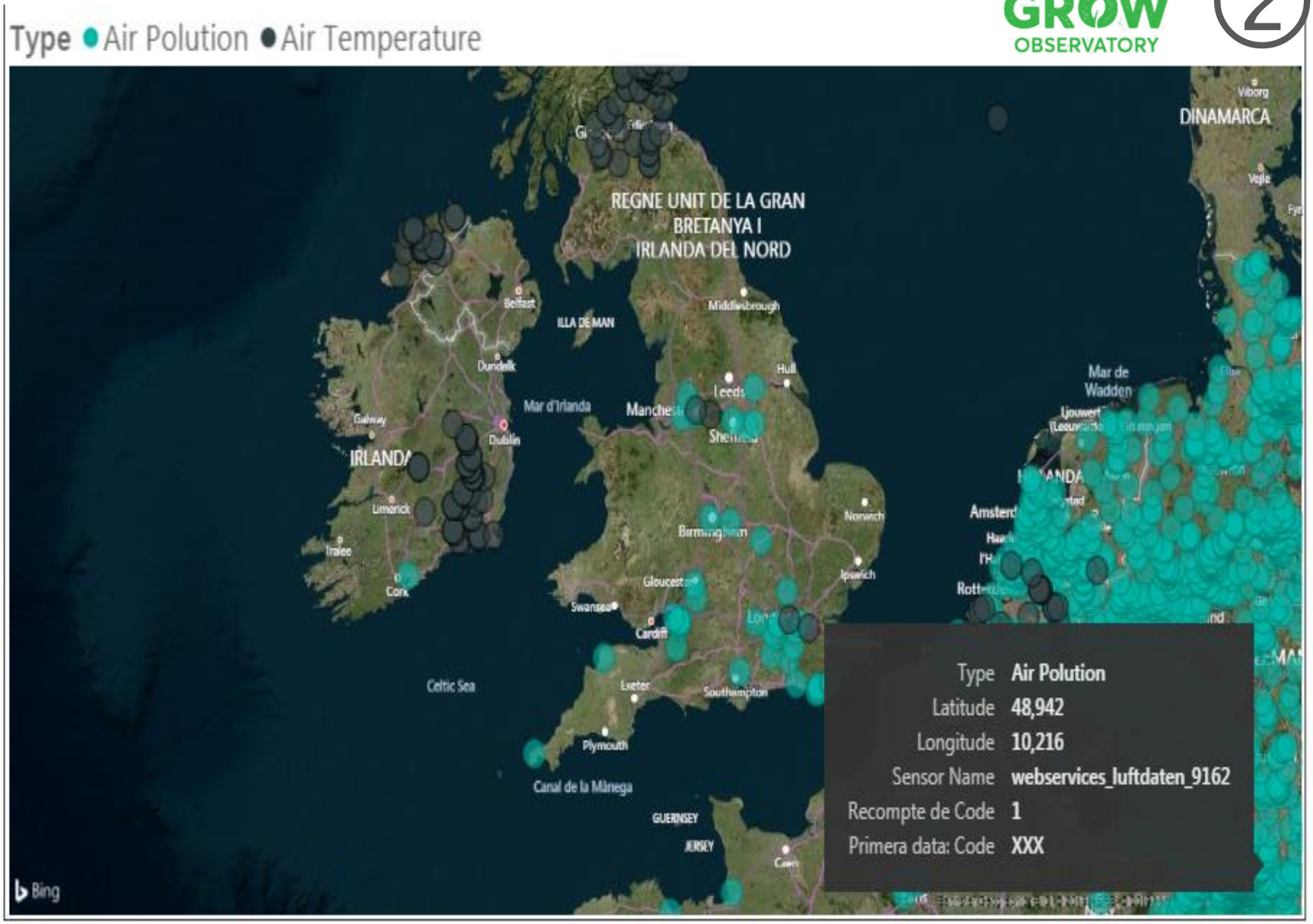
Architecture two

②

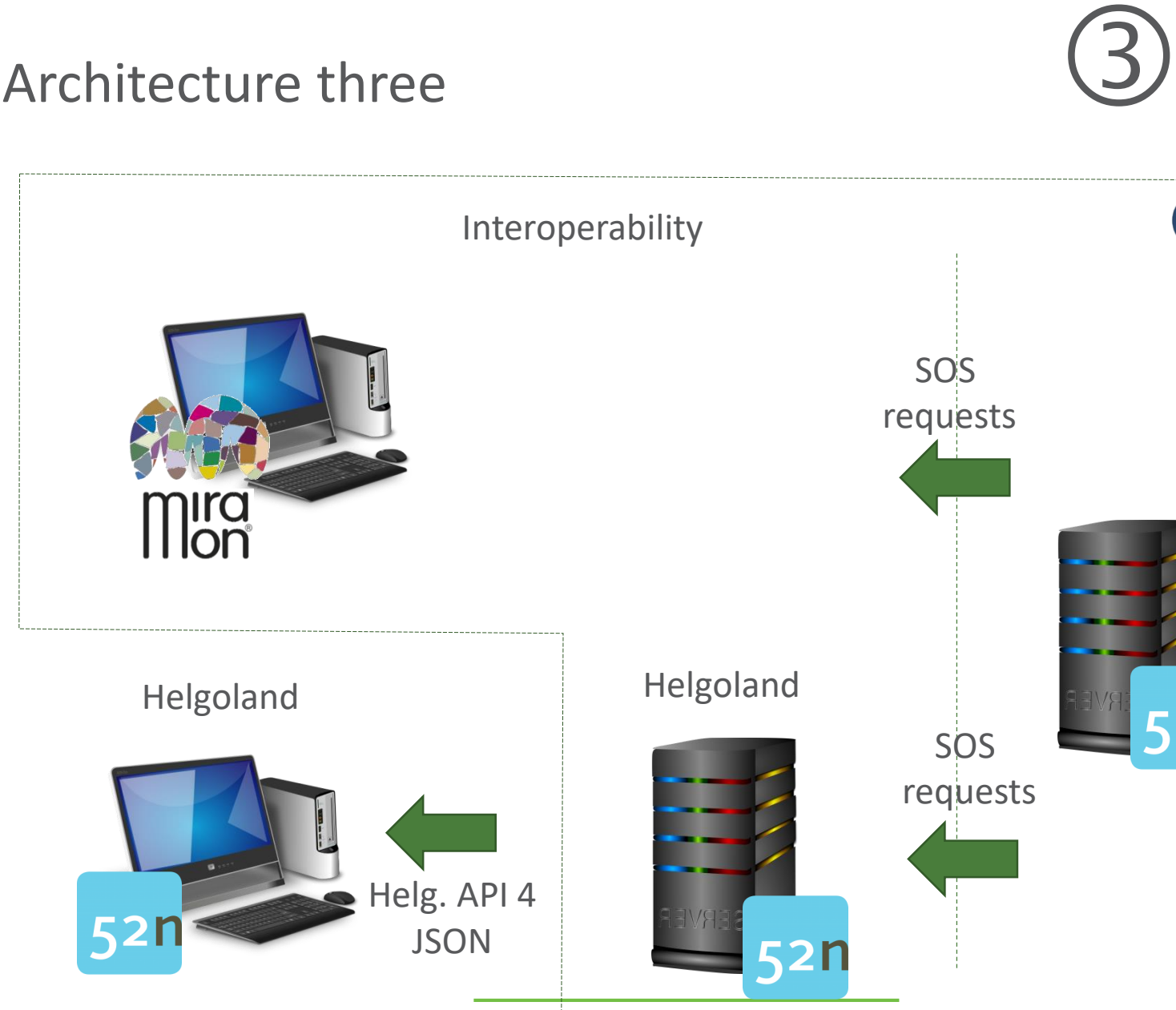
API 3



GROW OBSERVATORY

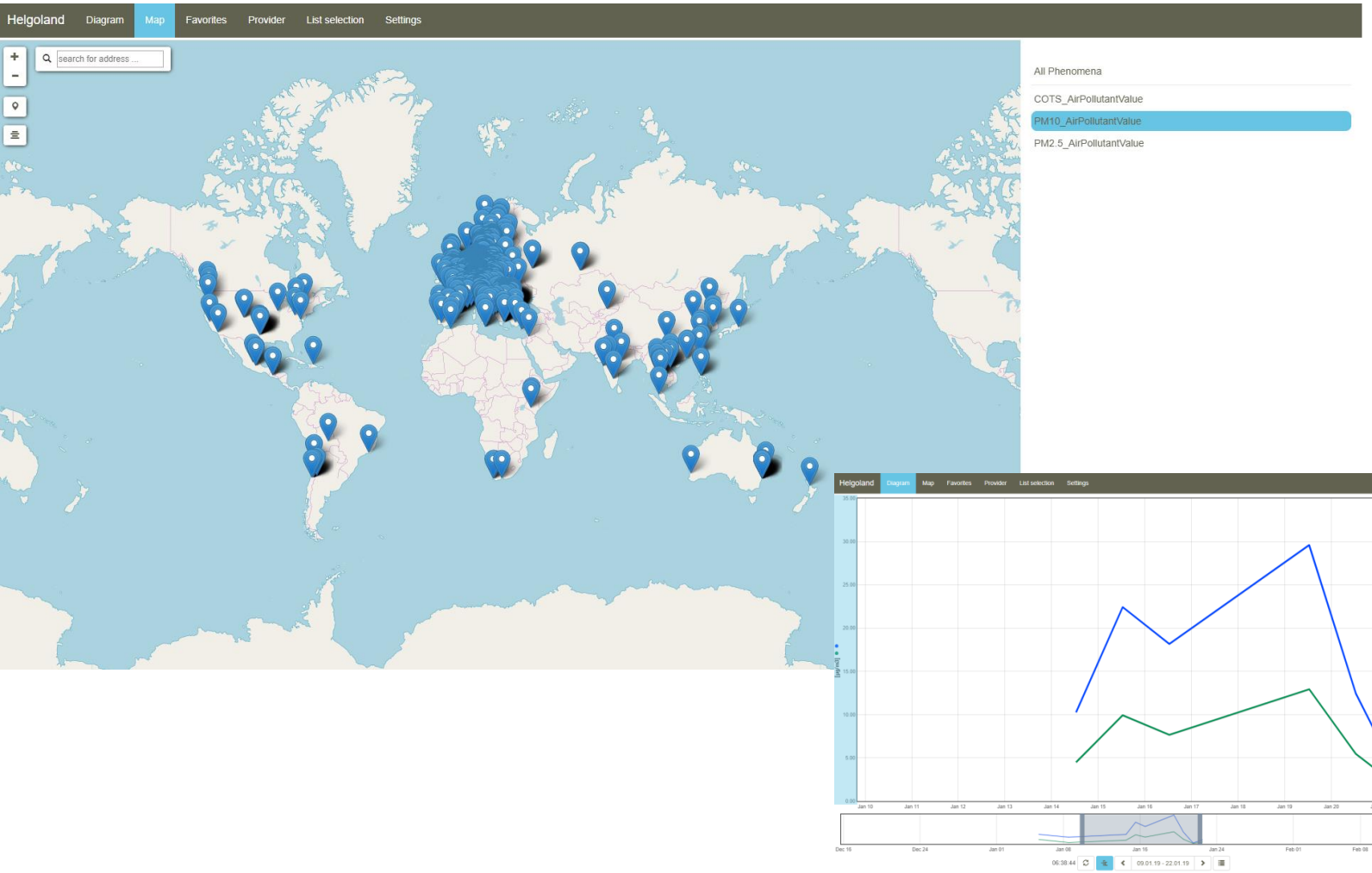


Architecture three



Helgolan SOS in action

③



How many APIs?

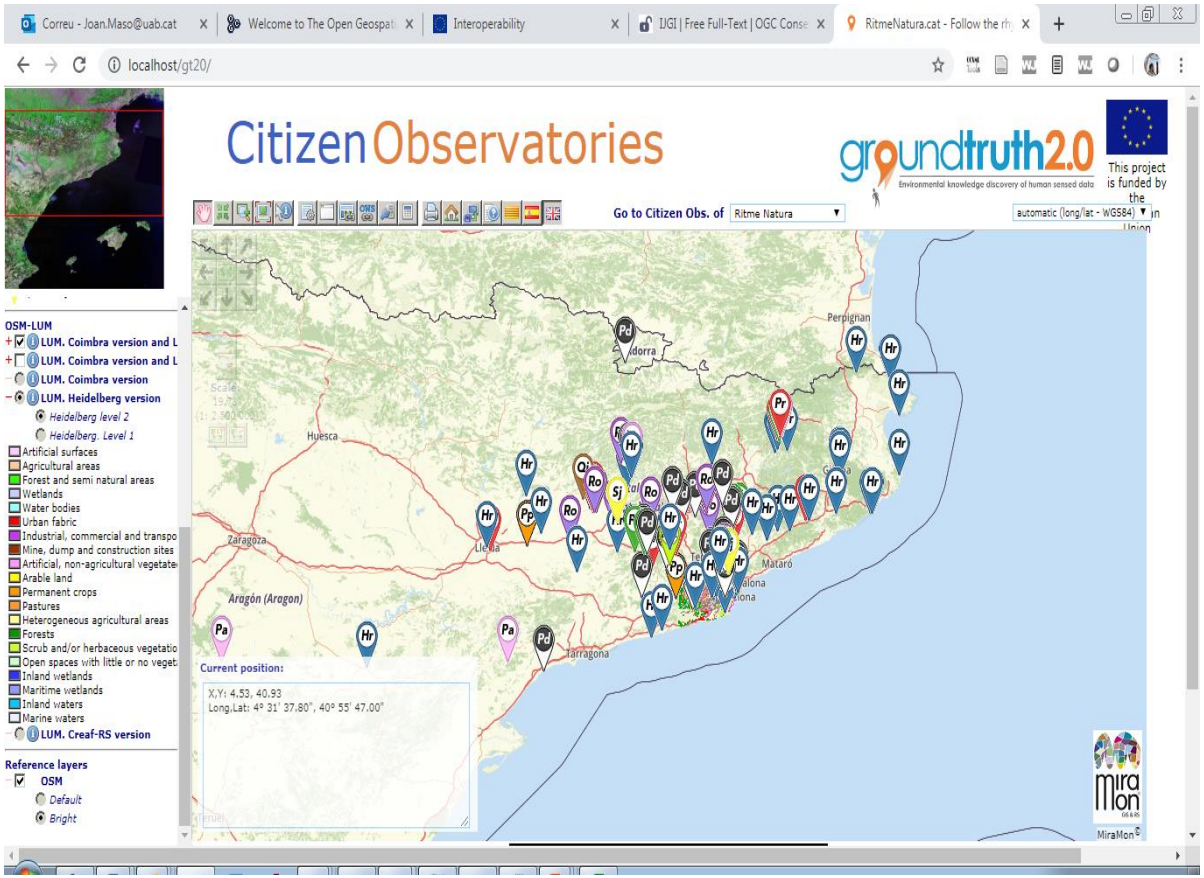
...but only one SOS

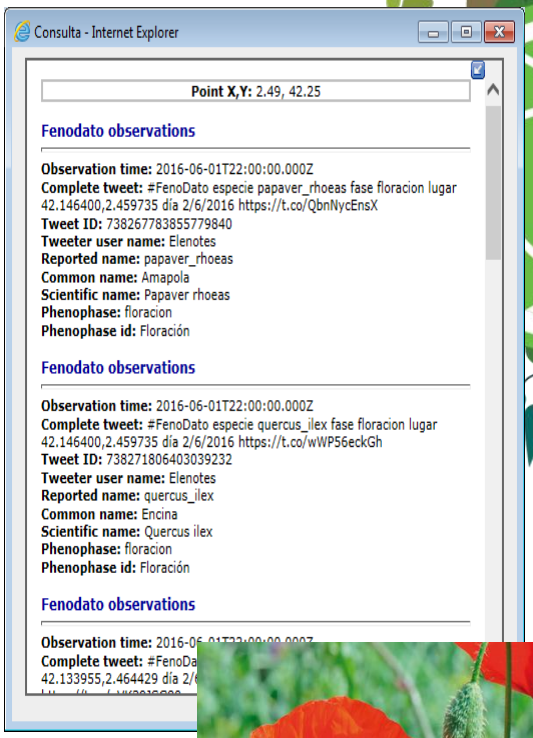
Details

Spanish case: Phenological observations exposed as WFS

- http://www.ogc3.uab.cat/cgi-bin/CitSci/MiraMon.cgi?VERSION=2.0.0&SERVICE=WFS&REQUEST=GetFeature&ATRIBUTFORMAT=complex&SRSNAME=EPSG:4326&TYPENAME=ritme-natura&OUTPUTFORMAT=application/json&BBOX=-1.2241759140625,40.22934409765625,4.7633729140625,43.10226890234375

groundtruth2.0
Environmental knowledge discovery of human sensed data





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


```
{
        "type": "FeatureCollection",
        "bbox": [40.2293440977, -1.2241759141, 43.1022689023, 4.7633729141],
        "features": [
        {
                "type": "Feature",
                "id": "ritme-natura.0",
                "bbox": [1.186299, 41.504417, 1.186299, 41.504417],
                "geometry": {
                        "type": "Point",
                        "coordinates": [1.1862990000, 41.5044170000]
                },
                "properties": {
                        "ID":760,
                        "observed_on_string":"17/04/2016",
                        "observed_on":"17/04/2016",
                        "time_observed_at":"",
                        "TIME_ZONE":"UTC",
                        "USER_ID":233,
                        "USER_LOGIN":"jaume-piera",
                        "CREATED_AT":"2016-04-19 15:25:27 UTC",
                        "UPDATED_AT":"2017-09-07 10:13:36 UTC",
                        "quality_grade":"research",
                        "LICENSE":"CC-BY-SA",
                        "URL":"http://natusfera.gbif.es/observations/760",

"IMAGE_URL":"https://natusfera.gbif.es/attachments/local_photos/files/581/medium/20160417_Delichon_urbica_Belltall.jpg?1461(
                        "TAG_LIST":"",
                        "description":"Oreneta cuablanca fent el niu a l'església de Belltall (Tarragona)#House Martin nests
church of Belltall (Tarragona)#Avión común haciendo el nido en la iglesia de Belltall (Tarragona)",
                        "ID_PLEASE":"false",
                        "num_identification_agreements":1,
                        "num_identification_disagreements":0,
                        "captive_cultivated":"false",
                        "place_guess":"Belltall",
```

Swedish case as SOS service

- http://www.ogc3.uab.cat/cgi-bin/CitSci/MiraMon.cgi?VERSION=2.0.0&SERVICE=SOS&REQUEST=GetObservation&featureOfInterest=http://www.opengis.uab.cat/vatten-fokus/featureOfInterest/&SRSNAME=EPSG:4326&BBOX=58.429753298828125,15.827924792968751,59.866215701171875,18.82169920703125

Citizen Observatories

groundtruth2

Environmental knowledge discovery of human sensed data

Go to Citizen Obs. of Vatten Fokus

automatic

OSM-LUM
- LUM. Coimbra version and LUM
- LUM. Coimbra version and LUM
- LUM. Coimbra version
- LUM. Heidelberg version
  - Heidelberg level 2
  - Heidelberg. Level 1
- LUM. Creaf-RS version
- Vatten Fokus
  - Nitrate
  - Phosphate
  - Water color

mg/L
- Clean waters. No risk of eutrophication
- Water at risk of eutrophication
- Polluted waters. Risk of strong eutrophication

Reference layers
- OSM
  - Default
  - Bright

Current position:
X,Y: 15.91, 59.78
Long,Lat: 15° 54' 47.05", 59° 46' 51.86"

Consulta - Google Chrome

localhost/gt20/consulta_de_cop.htm

Vatten Fokus

User ID:
Sample ID: 39544
Creation Date: 21/04/2018 17:55
Modification date: 21/04/2018 17:55
Sample date: 21/04/2018 0:00
Group ID: Dunkern, Group ID: 38438
Site name: Hemma
Sample date/time: 21/04/2018 0:00
Total number of participants: 4
Notes:
Freshwater body type: Other
Other freshwater body type: Lake
Land use in the immediate surroundings: Agriculture
Other the land use in the immediate surroundings:
Bank vegetation: Grass
Other bank vegetation:
On the water surface: LitterFoam
Pollution sources in the immediate surroundings: Other
Evidence of water uses: FishingBoatingPublic water supply
Other evidence of water uses:
Evidence of aquatic life: Plants below the surfaceFloating plantsFrogs/toads
Other evidence of aquatic life:
Algae presence: No algae
Estimated the water flow: Slow
Estimated water level: Average
Nitrate(mg/L): 1.50
Phosphate(mg/L): 0.075
Water Quality Secchi Tube (Turbidity):
Result:
Estimated water colour: Colourless
Other estimated water colour:

MiraMon

groundtruth2.0
Environmental knowledge discovery of human sensed data

Experiment 2: What is behind?

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

```
▼<sos:GetObservationResponse xmlns:sos="http://www.opengis.net/sos/2.0"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org/1999/xlink"
 xmlns:om="http://www.opengis.net/om/2.0" xmlns:gml="http://www.opengis.net/gml/3.2"
 xmlns:swe="http://www.opengis.net/swe/2.0" xsi:schemaLocation="http://www.opengis.net/sos/2.0
 http://schemas.opengis.net/sos/2.0/sosGetObservation.xsd http://www.opengis.net/gml/3.2
 http://schemas.opengis.net/gml/3.2.1/gml.xsd http://www.opengis.net/swe/2.0
 http://schemas.opengis.net/sweCommon/2.0/swe.xsd http://www.opengis.net/om/2.0
 http://schemas.opengis.net/om/2.0/observation.xsd">
  ▼<sos:observationData>
    ▼<om:OM_Observation gml:id="vatten-fokus_2_1">
       <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_ComplexObservation"/>
       <om:procedure xlink:href="http://www.opengis.uab.cat/vatten-fokus/procedure/22655"/>
       <om:observedProperty xlink:href="http://www.opengis.uab.cat/vatten-fokus/observedProperty"/>
       <om:featureOfInterest xlink:href="http://www.opengis.uab.cat/vatten-fokus/featureOfInterest/2"/>
      ▼<om:result xsi:type="swe:DataRecordPropertyType">
        ▼<swe:DataRecord>
          ▼<swe:field name="SAMPLE_ID">
            ▼<swe:Quantity definition="http://www.opengis.uab.cat/vatten-fokus/variable/SAMPLE_ID">
               <swe:uom/>
               <swe:value>45821</swe:value>
             </swe:Quantity>
           </swe:field>
          ▼<swe:field name="CREA_DATE">
            ▼<swe:Text definition="http://www.opengis.uab.cat/vatten-fokus/field/Creation_Date">
               <swe:value>07/12/2018 17:23</swe:value>
             </swe:Text>
           </swe:field>
          ▼<swe:field name="CHAN_DATE">
            ▼<swe:Text definition="http://www.opengis.uab.cat/vatten-fokus/field/Modification_date">
               <swe:value>07/12/2018 17:23</swe:value>
             </swe:Text>
           </swe:field>
          ▼<swe:field name="SAMPLEDATE">
            ▼<swe:Text definition="http://www.opengis.uab.cat/vatten-fokus/field/Sample_date">
               <swe:value>07/12/2018 15:00</swe:value>
             </swe:Text>
```

Belgium case as SOS service

- *http://www.ogc3.uab.cat/cgi-bin/CitSci/MiraMon.cgi?*VERSION=2.0.0&SERVICE=SOS&REQUEST=GetObservation&featureOfInterest=http://www.opengis.uab.cat/meet-mee-mechelen/featureOfInterest/&SRSNAME=EPSG:4326&BBOX=50.864041514648434,4.085620198242188,51.18951148535156,4.834063801757813&responseFormat=application/json

groundtruth2.0
Environmental knowledge discovery of human sensed data

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


```
{
	"request" : "GetObservation",
	"version" : "2.0.0",
	"service" : "SOS",
	"crs": {
		"type": "name",
		"properties": {
			"name": "EPSG:4326"
		}
	},
	"observations":[
		{
			"id":"meet-mee-mechelen_1_0",
			"type" : "http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_ComplexObservation",
			"phenomenonTime" : "2017-11-19 17:20:00+01",
			"resultTime" : "2017-11-19 17:20:00+01",
			"procedure" : "http://www.opengis.uab.cat/meet-mee-mechelen/procedure/5",
			"observedProperty" : "http://www.opengis.uab.cat/meet-mee-mechelen/observedProperty",
			"featureOfInterest" : "http://www.opengis.uab.cat/meet-mee-mechelen/featureOfInterest/1",
			"result": {
				"type":"DataRecord",
				"field":[
					{
						"name" : "CAMPAIGN",
						"type" : "Text",
						"definition" :"http://www.opengis.uab.cat/meet-mee-mechelen/field/CAMPAIGN",
						"value" : "Oct-Nov2017"
					},
					{
						"name" : "time_first",
						"type" : "Text",
						"definition" :"http://www.opengis.uab.cat/meet-mee-mechelen/field/time_first"
						"value" : "2017-11-06 08:00:10+01"
					},
					{
						"name" : "bc_aggr",
						"type" : "Quantity",
						"definition" :"http://www.opengis.uab.cat/meet-mee-mechelen/variable/bc_aggr"
						"value" : "3155"
```

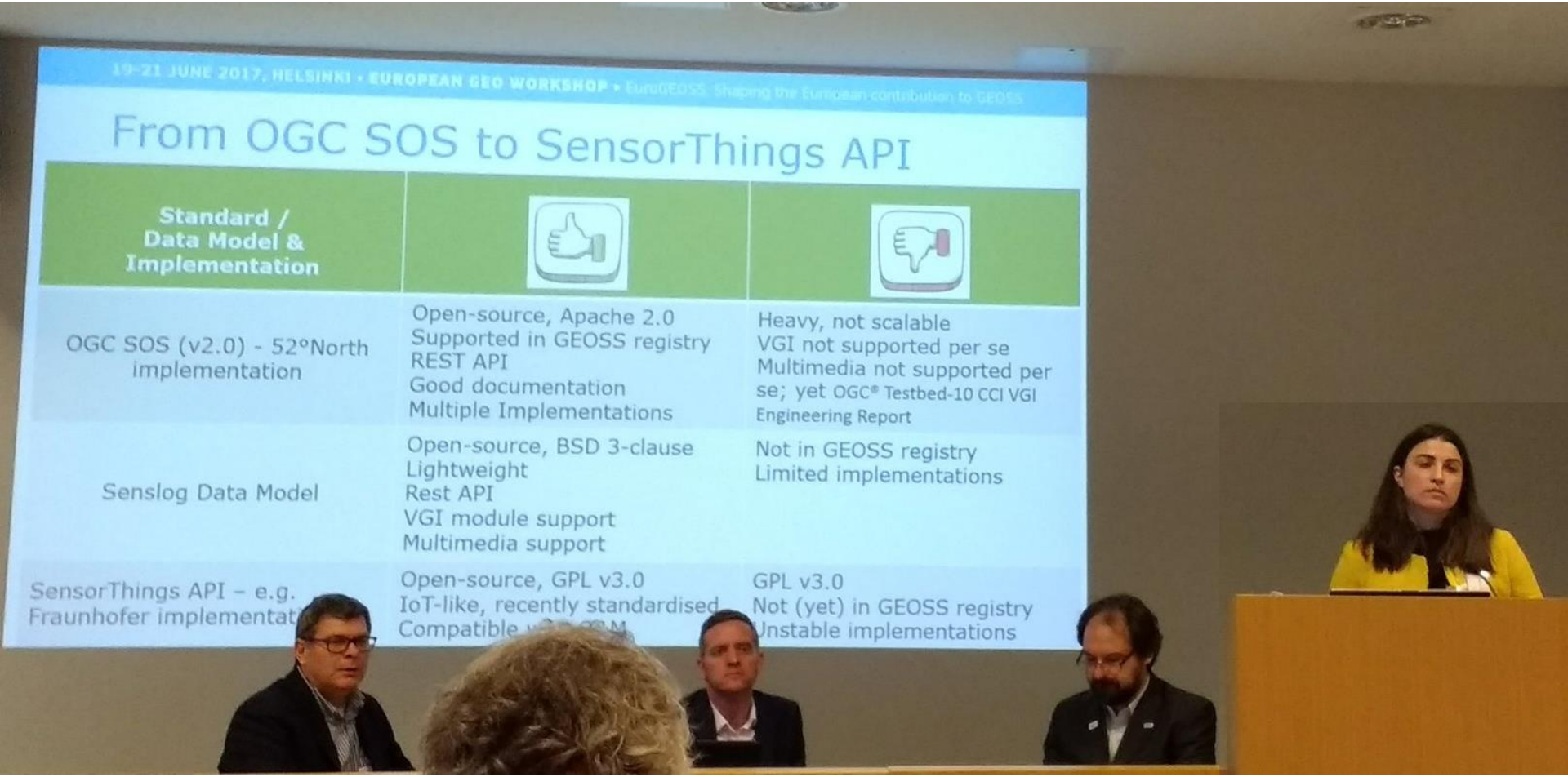
Some conclusions

- 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-JSON-”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



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)

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)

CitSciIE

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

Time Line

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

How can you help

- Nominate judges (August 2019)
- Nominate datasets (August 2019)
- Spread the word from October 2019

Contact

- aecobley@Dundee.ac.uk
- s.m.coulson@dundee.ac.uk

groundtruth2.0

Environmental knowledge discovery of human sensed data

THANK YOU!

Any Questions?

Joan Masó

CREAF

joan.maso@uab.cat

CREAF

This project has received funding from the EU's Horizon 2020 research and innovation programme under GA no 769926 and 689744