

EU Framework Program for Research and Innovation (SC5-18a-2014 - H2020)

ConnectinGO

Project Nr: 641538

**Coordinating an Observation Network of Networks EnCompassing saTellite and IN-situ
to fill the Gaps in European Observations**

Deliverable D3.5

***Results of the Data Management Plan (DMP) and Report on the
participation in the Pilot on Open Research Data in Horizon 2020***

Version 1

Due date of deliverable: 31/12/2016

Actual submission date: 24/03/2017

Document control page	
Title	D3.5 Results of the Data Management Plan (DMP) and Report on the participation in the Pilot on Open Research Data in Horizon 2020
Creator	CREAF
Editor	CREAF
Description	Revision and description of the results of the Data Management Plan (DMP) and participation in the Pilot on Open Research Data in Horizon 2020 as requested in http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf . The review will be conducted in collaboration with WP1 and WP5.
Publisher	ConnectinGEOConsortium
Contributors	ConnectinGEOPartners
Type	Text
Format	MS-Word
Language	EN-GB
Creation date	30/12/2016
Version number	1.0
Version date	24/03/2017
Last modified by	Joan Masó
Rights	Copyright © 2017, ConnectinGEOConsortium
Dissemination level	<input type="checkbox"/> CO (confidential, only for members of the consortium)
	<input checked="" type="checkbox"/> PU (public)
	<input type="checkbox"/> PP (restricted to other programme participants)
	<input type="checkbox"/> RE (restricted to a group specified by the consortium)
When restricted, access granted to:	
Nature	<input checked="" type="checkbox"/> R (report)
	<input type="checkbox"/> P (prototype)
	<input type="checkbox"/> D (demonstrator)
	<input type="checkbox"/> O (other)
Review status	<input type="checkbox"/> Draft <i>Where applicable:</i>
	<input type="checkbox"/> WP leader accepted Accepted by the PTB
	<input type="checkbox"/> PMB quality controlled Accepted by the PTB as public document
	<input checked="" type="checkbox"/> Coordinator accepted
Action requested	<input type="checkbox"/> to be revised by all ConnectinGEO partners
	<input type="checkbox"/> for approval of the WP leader
	<input type="checkbox"/> for approval of the PMB
	<input type="checkbox"/> for approval of the Project Coordinator
	<input type="checkbox"/> for approval of the PTB
Requested deadline	

Revision history			
Version	Date	Modified by	Comments
0.1	30-12-2016	IS_CREAF	Created the basic content of the deliverable
0.2	12-01-2017	IS_CREAF	Review and update of the content
0.3	10-02-2017	JM_CREAF	Addition of new items and content
0.4	15-03-2017	IS_CREAF	New datasets included
1.0	24-03-2017	JM_CREAF	Final edits

Contributors	
Acronym	Full name
IS_CREAF	Ivette Serral (CREAF)
JM_CREAF	Joan Masó (CREAF)

Copyright © 2017, ConnectinGEO Consortium

The ConnectinGEO Consortium grants third parties the right to use and distribute all or parts of this document, provided that the ConnectinGEO project and the document are properly referenced.

THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Table of Contents

1.	Introduction	1
2.	Review on the Data Management Plan report (D3.3)	1
2.1.	<i>Technical and support documentation about the use of the ConnectinGEO methodology</i>	1
2.1.1.	<i>Data set reference and name</i>	1
2.1.2.	<i>Data set description</i>	1
2.1.3.	<i>Results in standards and metadata</i>	2
2.1.4.	<i>Results in data sharing</i>	2
2.1.5.	<i>Results in archiving and preservation (including storage and backup)</i>	2
2.2.	<i>Observation gaps survey results</i>	2
2.2.1.	<i>Data set reference and name</i>	2
2.2.2.	<i>Data set description</i>	2
2.2.3.	<i>Results in standards and metadata</i>	3
2.2.4.	<i>Results in data sharing</i>	3
2.2.5.	<i>Results in archiving and preservation (including storage and backup)</i>	3
2.3.	<i>Observation Inventory for ConnectinGEO</i>	3
2.3.1.	<i>Data set reference and name</i>	3
2.3.2.	<i>Data set description</i>	3
2.3.3.	<i>Results in standards and metadata</i>	5
2.3.4.	<i>Results in data sharing</i>	5
2.3.5.	<i>Results in archiving and preservation (including storage and backup)</i>	6
2.4.	<i>Inventory of user needs from the Socio-Economic and Environmental Information Needs Knowledge Base (SEE IN KB)</i>	6
2.4.1.	<i>Data set reference and name</i>	6
2.4.2.	<i>Data set description</i>	6
2.4.3.	<i>Results in standards and metadata</i>	7
2.4.4.	<i>Results in data sharing</i>	8
2.4.5.	<i>Results in archiving and preservation (including storage and backup)</i>	8
2.5.	<i>Solar renewable energy industrial challenge data</i>	8
2.5.1.	<i>Data set reference and name</i>	8
2.5.2.	<i>Data set description</i>	8
2.5.3.	<i>Results in standards and metadata</i>	9
2.5.4.	<i>Results in data sharing</i>	10
2.5.5.	<i>Results in archiving and preservation (including storage and backup)</i>	10
2.6.	<i>ENEON Commons</i>	10
2.6.1.	<i>Data set reference and name</i>	10
2.6.2.	<i>Data set description</i>	10
2.6.3.	<i>Results in standards and metadata</i>	10
2.6.4.	<i>Results in data sharing</i>	11

2.6.5.	<i>Results in archiving and preservation (including storage and backup)</i>	11
2.7.	<i>ENEON composition and EO network graph</i>	11
2.7.1.	<i>Data set reference and name</i>	11
2.7.2.	<i>Data set description</i>	11
2.7.3.	<i>Results in standards and metadata</i>	12
2.7.4.	<i>Results in data sharing</i>	13
2.7.5.	<i>Results in archiving and preservation (including storage and backup)</i>	14
2.8.	<i>ENEON's graph user feedback</i>	14
2.8.1.	<i>Data set reference and name</i>	14
2.8.2.	<i>Data set description</i>	15
2.8.3.	<i>Results in standards and metadata</i>	15
2.8.4.	<i>Results in data sharing</i>	15
2.8.5.	<i>Results in archiving and preservation (including storage and backup)</i>	15
2.9.	<i>Observation gaps table</i>	15
2.9.1.	<i>Data set reference and name</i>	15
2.9.2.	<i>Data set description</i>	15
2.9.3.	<i>Results in standards and metadata</i>	17
2.9.4.	<i>Results in data sharing</i>	17
2.9.5.	<i>Results in archiving and preservation (including storage and backup)</i>	17
2.10.	<i>Observation gaps user feedback</i>	18
2.10.1.	<i>Data set reference and name</i>	18
2.10.2.	<i>Data set description</i>	18
2.10.3.	<i>Results in standards and metadata</i>	18
2.10.4.	<i>Results in data sharing</i>	18
2.10.5.	<i>Results in archiving and preservation (including storage and backup)</i>	18
2.11.	<i>Information regarding the Essential Variables</i>	18
2.11.1.	<i>Data set reference and name</i>	18
2.11.2.	<i>Data set description</i>	18
2.11.3.	<i>Results in standards and metadata</i>	18
2.11.4.	<i>Results in data sharing</i>	18
2.11.5.	<i>Results in archiving and preservation (including storage and backup)</i>	19
3.	<i>Report on the participation in the Pilot on Open Research Data in Horizon 2020</i>	19

1. Introduction

ConnectinGEO's main mission is not the creation of geospatial datasets but establishing an inventory of data and observations delivered by existing EO networks in Europe. ConnectinGEO also collects relevant standards and the contribution to the GCI standards registry. However, some data can be a result from the Industry challenges.

Having all this information on data, observations and standards, ConnectinGEO will apply a methodology for detecting gaps and redundancies. The result of this methodology is also a data output of the project and will be collected and openly offered as well.

The purpose of the Data Management Plan (DMP) is to help in considering any problems or challenges that may be encountered during the project and in identifying ways to overcome these. A DMP should be thought of as a "living" document outlining how the research data collected or generated will be handled during and after a research project.

The Data Management Plan developed at the beginning of the project (D3.3), has been followed during ConnectinGEO's life and it is being updated on this deliverable.

This DMP takes into consideration the following datasets:

1. Technical and support documentation about the use of the ConnectinGEO methodology
2. Observation gaps survey results
3. Observation Inventory for ConnectinGEO
4. Inventory of user needs from the Socio-Economic and Environmental Information Knowledge Base (SEE IN KB)
5. Solar renewable energy industrial challenge data
6. ENEON Commons
7. ENEON composition and EO network graph
8. ENEON's graph user feedback
9. Observation gaps table
10. Observation gaps user feedback
11. Information regarding the Essential Variables

2. Review on the Data Management Plan report (D3.3)

2.1. Technical and support documentation about the use of the ConnectinGEO methodology

2.1.1. Data set reference and name

Technical and support documentation to use of the ConnectinGEO methodology.

2.1.2. Data set description

The ConnectinGEO methodology for identifying the gaps and assessing prioritisation has been published in ConnectinGEO website and in *D6.1 ConnectinGEO methodology* (available at www.connectingeo.net and the project digital repository: <http://ddd.uab.cat/collection/connectingeo?ln=en>).

The methodology covers five approaches that ConnectinGEO executed:

1. Top-Down thread 1: Identification of a **collection of observation requirements and specifications** from **generic goals** for sustainability of the global civilization as expressed in the GEOSS Strategic Targets, the SDGs, and the adherence to the planetary boundaries.
2. Top-Down thread 2: Incorporation of material from international programs such as Future Earth, Belmont Forum, the Research Data Alliance, and community assessments of socio-economic benefits of Earth observations.
3. Bottom-up thread 1: A **consultation process** in the current EO networks, consisting of collaboration platforms, surveys and discussions at workshops and even involvement of citizen science.
4. Bottom-up thread 2: A careful **analysis of the observations and measurements** that are **currently** available in GEOSS Discovery and Access Broker complemented by other means (e.g. scientific literature).
5. Bottom-up thread 3: The realization of a series of real **industry-driven challenges** to assess the problems and gaps emerging during the creation of business opportunities.

The deliverable also covers the results of applying the methodology into GEOSS.

2.1.3. Results in standards and metadata

No standards and metadata apply to this issue.

2.1.4. Results in data sharing

The webpage is freely accessible.

2.1.5. Results in archiving and preservation (including storage and backup)

The webpage is maintained and preserved by CREAM.

2.2. Observation gaps survey results

2.2.1. Data set reference and name

Observation gaps survey results.

2.2.2. Data set description

ConnectinGEO aimed to facilitate a broader and more accessible knowledge base to support the needs of the GEO SBAs and their users.

To better detect gaps in the accessibility of observation data, an online survey to collect feedback from scientists and potential users was developed. The idea of this survey was to help identify existing issues with data availability, technology, data policies and data quality which needed to be addressed to improve and support the work in different SBAs.

This survey was created and conducted using Google Forms. The results of this survey have been exported for further processing while the original survey was removed from the server.

The survey is still available via this URL: <http://goo.gl/forms/ql7JvS6id0>

2.2.3. Results in standards and metadata

The data has been stored by Google Forms in a simple tabular form. Thus, no specific standards were used. However, the export of the data was performed using a CSV (comma separated values) structure.

2.2.4. Results in data sharing

The results of the survey were published in an aggregated form as part of the ConnectinGEO *D6.2 Gap analysis draft report*. The survey was structured with a low entry barrier to ensure a higher number of responses; the participants were not requested to allow the publication of their personal information associated to their opinions. Thus, to preserve private rights on personal data, the raw data of the survey was not published and is only accessible by 52N for the purposes of the survey. Certain aggregated results of the query can be made accessible in anonymous form.

2.2.5. Results in archiving and preservation (including storage and backup)

The deliverable (D6.2) has been stored on the web site and the project digital repository. The raw data with the response of the survey will be preserved by 52N that will ensure that it will not be accessible.

2.3. Observation Inventory for ConnectinGEO

2.3.1. Data set reference and name

ConnectinGEO observation inventory.

2.3.2. Data set description

In the initial steps of the project, an inventory of observations data sources was created, based on GCI Information, the DAB, and Copernicus services catalogues that allowed to assess the coherency, continuity, accuracy and accessibility of the information available. It provided an assessment of the availability and accessibility of data required to derive the information (e.g. EV) or to run the models that generate the forecasts and projections, and it helped to support the prioritization of the gaps that need to be addressed.

The information and parameters gathered and queried for the observation inventory have been:

- spatial, temporal, and thematic coverage (considering missing data)
- accuracy/quality (metadata as well as data)
- accessibility (including licenses, standards, and formats)
- cost aspects (e.g. GEOSS Data CORE, fees)
- versions of observation descriptions.

The ConnectinGEO observation inventory data was complemented with previous inventories analysis and Copernicus services catalogues. A form of access through the DAB was analysed and implemented. Efforts were undertaken to equally consider remote

sensing, airborne and in-situ sensors. The new capabilities, introduced by Sentinel missions, have been considered.

The observation inventory was completed by the description of the observations and measurements automatically (for example, including sensor specifications or hierarchical relations) and with information extracted from scientific literature.

The following list presents the catalogues in the DAB and its number of records:

New Zealand government geodata catalogue	2175
ISPRA Monitoring network	7680
Chile Geoportal	11647
South African Environmental Observation Network	14818
WIS GIS DWD	143879
IGN GeoPortal	50
UK Data Gov (under testing)	6089
FedEO	2551
New Zealand Monitoring Network	32
NIWA data catalogue	188
Geoscience Australia	21301
Red Vial (Road Network, from Ministry of Public Works) -- from Chile	16
HIS Central US	14331907
NOAA Unified Access Framework Catalogue	5114
Limites (Administrative Boundaries, from Ministry of Public Works as well) -- from Chile	5
SeaDataNet	476
EEA SDI Catalogue	414
GMOS Database	819
Global River Discharge Datasets (GRDC/GEOWOW) - Kisters AG	1928
Registered Data	44417
IRIS Event	4163124
IRIS Station	484768
Data Integration and Analysis System (DIAS) - Japan	217
NASA Global Change Master Directory	28108
RASAQM	297
EGASKRO	408
Knossos	536
INPE CDSR	863290
ArcGIS Online ESRI	185000
BYU World Water Data catalog	15
IODE	7509
GBIF	19888012
MEDINA SDI	158
Canadian Monitoring Network	828
CEOS WGISS Integrated Catalogue (CWIC)	1852
Webservice Energy Catalogue	1699
PANGAEA	335877
US Data Gov	85229
One Geology	438

The total number of record is: 40642371.

2.3.3. Results in standards and metadata

Metadata on accuracy/quality have been collected about the data included in the observation inventory. It has been originated from current information in the metadata concentrated in the DAB and the GCI.

Metadata explaining the network source of the observations has also been also collected, to the following fields (mainly extracted from the ISO 19115 core metadata):

- Id
- Title [*Name of the network*]
- Abstract
- Producer
- Topic category (1..*)
- Theme/topic (0..*) [*SBA main domain*]
- Measurement (0..*)
- Keywords (0..*)
- EV (0..*)
- Spatial extent (0..1) [*Spatial scale of the action of the network*]
- Temporal extent (0..1) [*Temporal scale of the action of the network*]
- Resolution/scale (0..1)
- Series/parent resource (0..1)
- Link to data (0..*)

Extra extracted requirements are:

- Location of the action of the network
- Administrative aggregation level
- Date of creation
- Availability of background information (report, raw data, metadata)
- Concrete tools / data / guidelines published
- Comments

2.3.4. Results in data sharing

The data collected in the observation inventory has been integrated on the project. This database has feed the ConnectinGEO methodology on gap and redundancies analysis.

The access to this information is made public via a project website: <http://oi.geodab.eu/oi-client-beta/home/index-dev.html#>. No registration is required to access the database of the inventory. The information is available for consulting.

The observation inventory requirements, the database schema and the queryable fields, are also publicly available via the same webpage for the inventory users as a way to better understand its database and the design.

The database has been processed by the ConnectinGEO team and statistical analysis of the observation inventory database has been elaborated and made public with the actual data.

2.3.5. Results in archiving and preservation (including storage and backup)

The long-term preservation of this dataset is strongly related to the long-term preservation of the inventory. The inventory is being preserved by CNR, the DAB responsible party. If it's found useful, the inventory can be updated in the future following the same project methodology.

There is no preservation strategy for the DAB content due to the regular and constant update of the DAB. This way of preserving the inventory as a snapshot of the DAB current status is important. After the end of the project, it is reasonable to assume that the preservation of the inventory will be related with the persistence of the DAB as a key component for the GEOSS infrastructure. The cost and efforts of preserving the database will be assumed by the CNR.

2.4. Inventory of user needs from the Socio-Economic and Environmental Information Needs Knowledge Base (SEE IN KB)

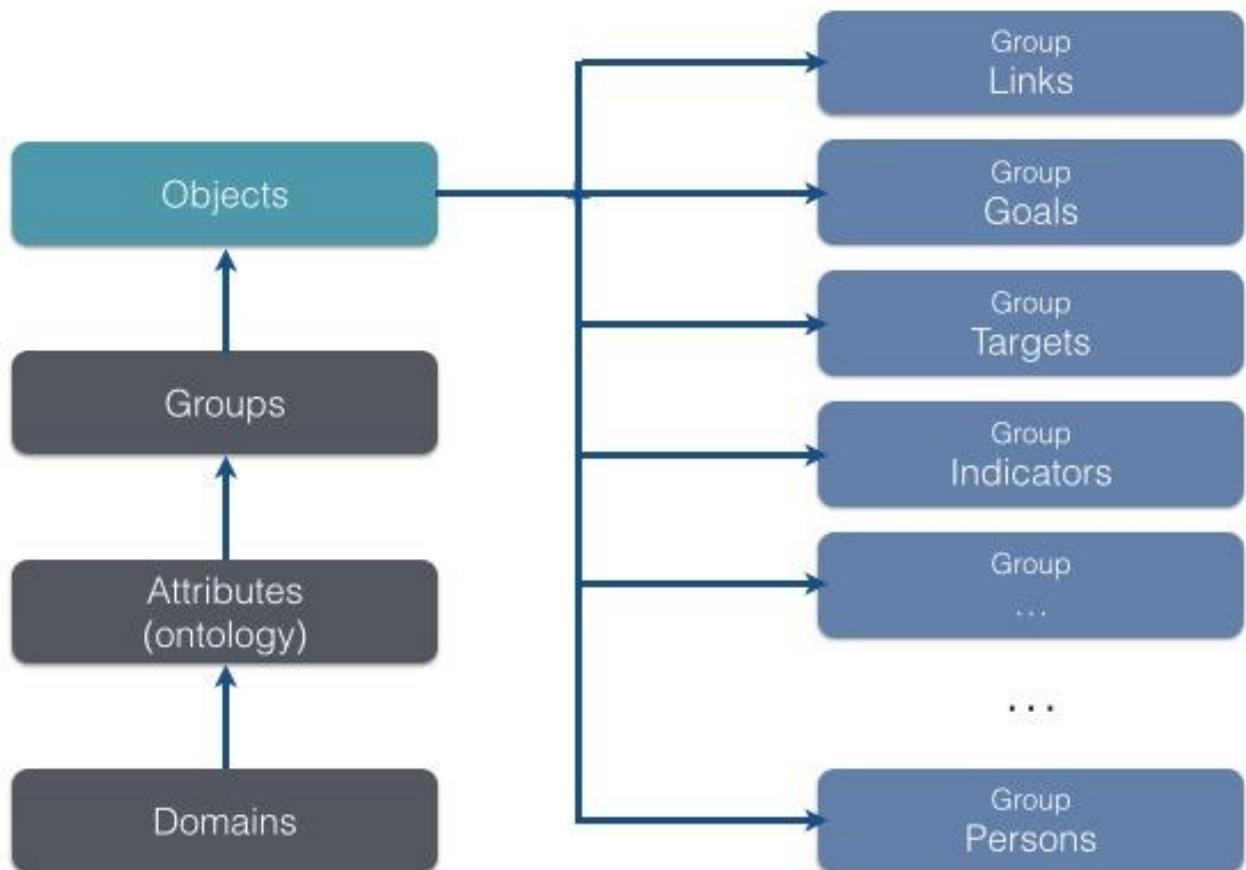
2.4.1. Data set reference and name

Inventory of user needs from Socio-Economic and Environmental Information Needs Knowledge Base (SEE IN KB).

2.4.2. Data set description

The GEOSS User Requirements Registry (URR) was the primary place to collect information on Earth observations requirements and the application and users that depended on these observations. A Ministerial guidance provided to GEO during the Ministerial Summit in 2014 requested that GEO “develop a comprehensive interdisciplinary knowledge base defining and documenting observations needed for all disciplines and facilitate availability and accessibility of these observations to user communities.” In response to this guidance, the URR has been transformed and extended into the SEE-IN KB.

The contents of the SEE IN KB reflect the knowledge, information, and data needs of a wide range of applications and users. The data model of the SEE-IN KB is based on unstructured objects (Figure 1). Each entity in the SEE-IN KB is an object. Objects can belong to one or more groups (for examples, see next paragraph). A group is defined by a set of attributes selected from a master set of attributes (the ontology). Examples of attributes are Variable, Resolution, Latency, Accuracy, etc. Each attribute comes with a specified domain taken from a master set of domains. Examples of domains are an interval of real number, a list of variables, or a list of specific terms. Groups also are associated with rules for how members of the group can act and interact with other objects. A special group are Links (which are also objects). Links define the relationship between two objects in the same or two different groups. This link concept is used to capture connectivity. With the group concept, new groups can be introduced as needed.



Main elements of the data model of the SEE-IN KB. The dark-gray boxes define the domains for attributes, the ontology, and the known groups. Each entity in the SEE-IN KB is an object, which are registered in an inventory, where the entity is associated with one or more groups. Details on the object are provided in the body of object, depending on the group membership. The objects in the Link group capture the connectivity and relationships between object.

The groups inherited from the URR include User Types, Applications, Requirements, Research Needs, Infrastructure Needs, Technology Needs, Capacity Needs, and Links. Additional groups include, among others, Persons, Models, Services, Datasets, Essential Variables, Variables, and several groups for various types of gaps.

The concept of Essential Variables (EVs) has been implemented into the SEE-IN KB through ConnectinGEO. Based on the working in ConnectinGEO WP2, rules have been made available to link societal goals (such as the Sustainable Development Goals, SDGs) to EVs. Existing set of EVs can be linked to societal goals and benefits. The results of the review of EV developments in the GEO SBAs have been published in the SEE-IN KB.

The SEE-IN KB has served as the primary source for “user needs” in ConnectinGEO and has been utilized for at least one of the gap analysis approaches in the ConnectinGEO methodology. Moreover, the outcomes of the ConnectinGEO gap analysis have been published in the SEE-IN KB. The information in the SEE-IN KB is not part of the Observation inventory but has been used on the project to compare offerings (data) with needs and determine the gaps. Also it has been used to set priorities.

2.4.3. Results in standards and metadata

Catalogue and metadata standards have been applied to the SEE-IN KB.

2.4.4. Results in data sharing

The SEE-IN KB is open access to the world-wide community. Guests can view and search all contents. Users can sign up and request editing authority. The SEE-IN KB has several options for user to publish objects in all available groups, establish new groups and develop rules for queries and analyses.

2.4.5. Results in archiving and preservation (including storage and backup)

The SEE-IN KB is developed as open-source software with community participation. It is a constantly evolving knowledge base and both information contents and rules can be updated through web-based interfaces. The gap analysis has been conducted by ConnectinGEO in real time.

2.5. Solar renewable energy industrial challenge data

2.5.1. Data set reference and name

Solar renewable energy industrial challenge data.

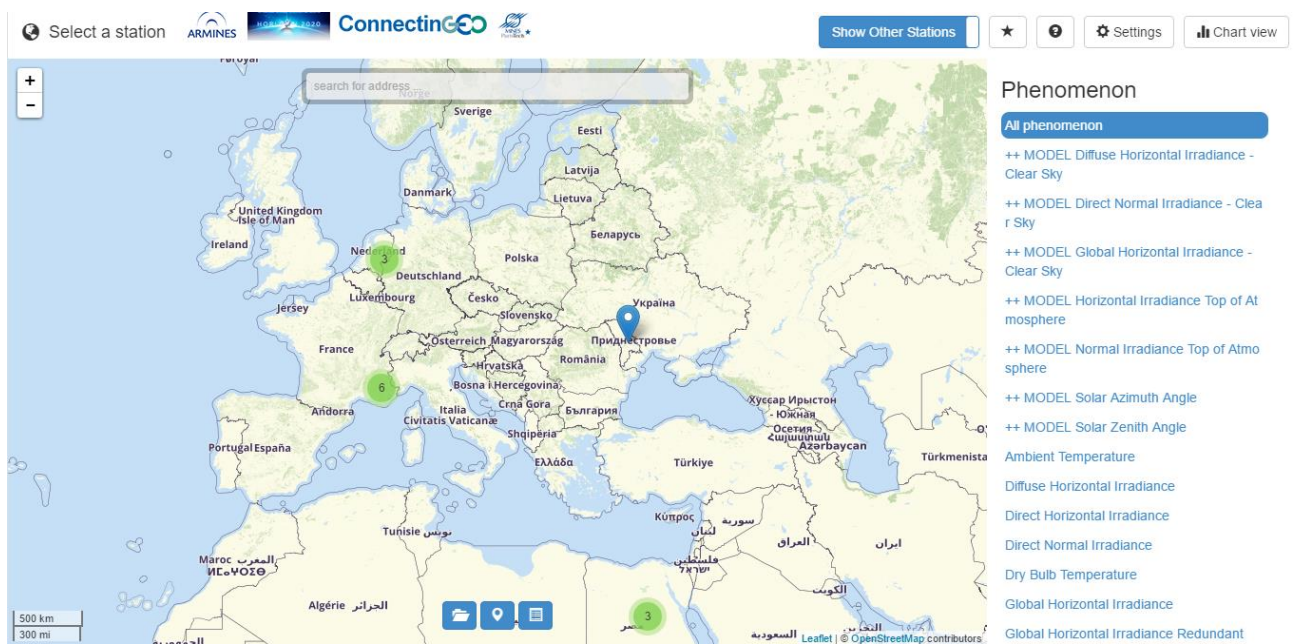
2.5.2. Data set description

Decision makers and project developers of solar power plants (photovoltaic, solar concentrating technologies) have identified the Surface Solar Irradiance (SSI) and its components as an EV for their business development. SSI observations are crucial in the process of selecting the most suitable location of new plants but there is only a sparse network of in-situ pyranometric stations. In the first phase satellite or numeric weather model data are used as a global location identification solution. In a second phase in-situ sensors are deployed for periods of at least one year to collect data, which are more detailed and can be the base for solid business models, which are crucial to acquire funding.

Nevertheless networks of in-situ pyranometric sensors, satellites, or numeric weather models are not able to fully answer the need of industrials and project developers if they are taken separately, but solutions can be found in the conjunction of these EO systems. Still missing is the means to exchange in-situ measurements across companies and between stakeholders in the market which could considerably improve business opportunities. In this task we will provide an open solution for interoperable exchange of SSI data comprising in-situ time-series observations as well as sensor descriptions. The practical solution will be based on the experience from other domains.

It will allow a comprehensive understanding of gaps in capacity and generally, more important, bridge the industrial and the research communities in GEO.

A dedicated platform has been developed in ConnectinGEO by 52N with pre-existing in-situ measurements for test but also with the important possibility for the companies to share their own legacy in-situ measurements to be part of this use case: <http://insitu.webservice-energy.org/jsClient-0.2.0>



WebGIS Client providing access to in-situ measurements

2.5.3. Results in standards and metadata

The platform has been integrated as a new component into an existing Spatial Data Infrastructure (SDI) dedicated to Energy and Environment namely the already existing and active Webservice-Energy.org GEOSS Community Portal. A key component of this Community Portal is the OGC Catalogue providing access to more than one thousand energy and environment datasets through interoperable Web-services. This catalogue is already registered in the GEOSS Common Infrastructure (GCI) and is harvested on a weekly basis by the GEOSS Discovery and Access Broker (DAB). Consequently the content of the catalogue is available for GEOSS user on the Geo Web Portal (GWP). The in-situ measurements deployed as form of sensor Web-services have been registered as metadata in the Webservice-Energy catalogue. Thus, the full scope from search and discovery of existing resources covered by the catalogue to the selection, display and download of measurement of interest from the dedicated sensor platform have been taken into account. The ability to provide the users with measurement download is a key aspect to start engaging the Energy community to share, release and use in-situ measurements. Participants provided their feedback on the benefit of such a platform and on the networking of resources for improving their daily work and increasing their business. Therefore this challenge has been a great tool to measure the impact of the ConnectinGEO methodology approach in a specific community and use case.

Regarding gaps in data models, standardized data and metadata models have been designed suitable for in-situ SSI observations as profiles of OGC Sensor Web Enablement standards. Profiles for XML encodings were defined as Schematron rules. The data and metadata models must take into account spatiotemporal coverage, lineage, data quality, IPR, the existing terminology and units of measure.

Regarding gaps in in-situ measurements, the catalogue platform with data management functions were extend based on the new (meta)data profiles and the GEOSS recommendations for interoperability based on 52°North Sensor Web components. The platform allows to

- (i) visualise sensor locations on a map,

- (ii) visualise measurements as time series plots and in tabular form,
- (iii) display sensor metadata at different levels of detail,
- (iv) upload of data by data owners, and
- (v) download raw observation data for offline processing. It will support company to prepare, pre-process and integrate their datasets.

2.5.4.Results in data sharing

An extension of the existing portal for the energy community allows access to data (particularly in-situ) instead of merely metadata and to allow data sharing between companies and organisations. This has allow to test the capacity of cooperation in the SSI community by introducing an unprecedented level of collaboration and eventually help to detect gaps in European EO networks. The pilot deployment has helped to set priorities in the energy domain. It has also explored the possibilities to work together with weather/climate networks.

2.5.5.Results in archiving and preservation (including storage and backup)

52N is the institution to store and preserve the data and the server/portal after the project end.

2.6. ENEON Commons

2.6.1.Data set reference and name

ENEON Commons database

2.6.2.Data set description

ENEON Commons is a collaborative platform where users can discuss, comment and submit queries about multidisciplinary problems. ENEON commons stores users interactions in the database as well as the proposed solutions to the questions formulated by users.

ENEON Commons is developed by 52N based on a marketplace based on the Open Source CMS Drupal that was developed for the WaterInnEU project. Within the ConnectinGEO project, this development was used as basis for building a first prototype of an ENEON Commons portal.

ENEON Commons is a collaboration environment offered through a Web portal (SaaS) developed and maintained by 52N. It offers a forum for debates and exchange of information, as well as, a queryable graph of networks and their observation types connected to the actual data through the GEOSS Discovery and Access Broker GEO-DAB. It accepts questions that are redirected to the appropriate network experts as well as providing automatically generated results (based on previous answers, user feedback, the GEODAB and the GEOSS Knowledge Base). It is also a Web service (PaaS) that can be integrated in other systems.

2.6.3.Results in standards and metadata

ENEON stores the data in a database. All interactions are timestamp. No metadata standards has been applied.

2.6.4. Results in data sharing

ENEON commons will make available information to help users to find the right Observation system that is able to provide the data they need:

- In-situ observation networks are introduced, exposed and made discoverable,
- users (e.g. scientists) can search for and request data for interdisciplinary problems,
- users can establish contacts with those networks that may deliver needed data,
- participants can informally share information, etc.

The information stored in the ENEON Commons is exposed as a graph in JSON and RDF. No personal data of the users is disclosed.

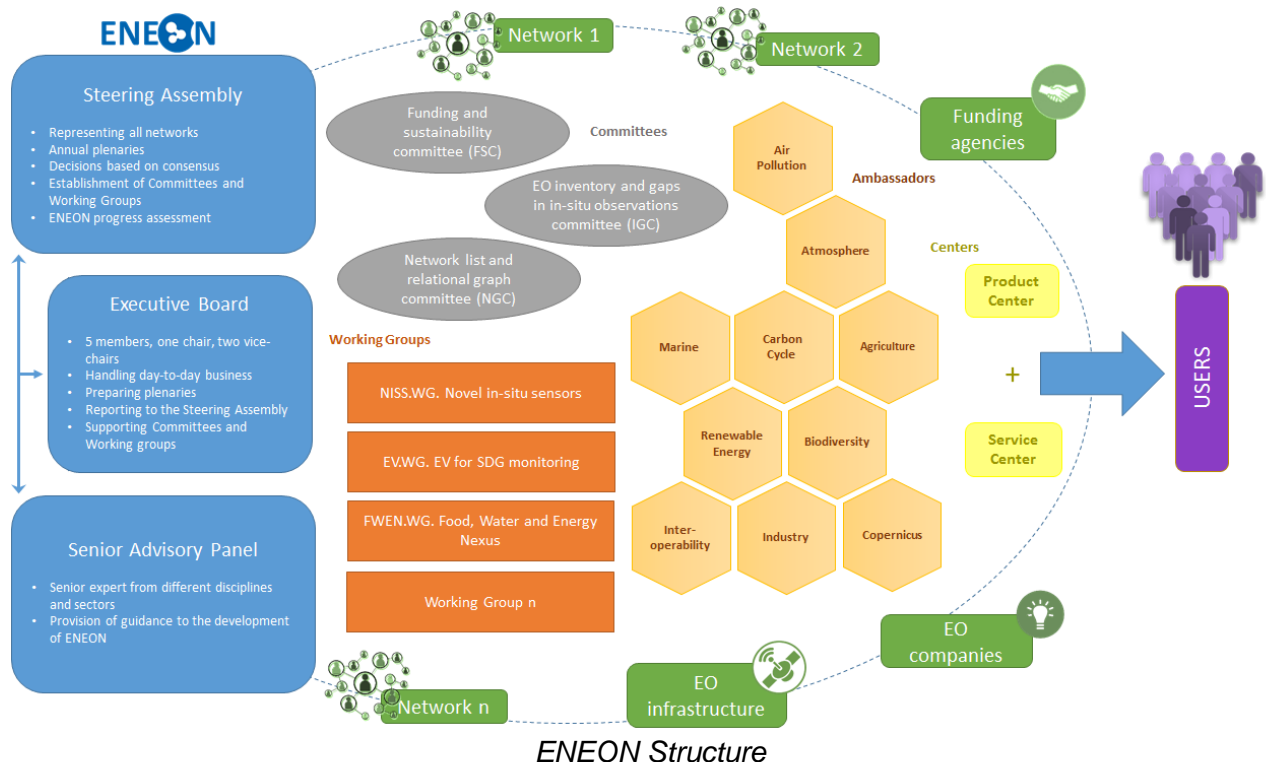
2.6.5. Results in archiving and preservation (including storage and backup)

52N is the institution to store and preserve the data and the server/portal after the project end. 52N has made available the code of the ENEON for future reuse.

2.7. ENEON composition and EO network graph

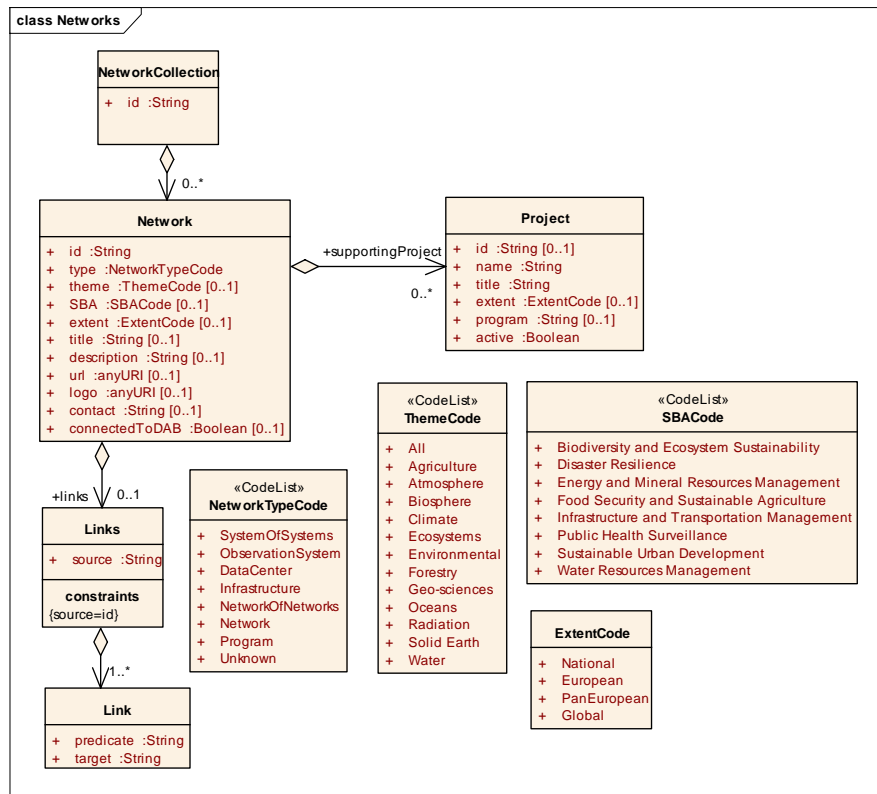
2.7.1. Data set reference and name

ENEON composition and EO network graph.



2.7.2. Data set description

Each individual network is defined by a reduced set of properties, including links to other networks as well and specifying the role of the relation. The ENEON graph follows a data model illustrated in the following figure:



UML class diagram for the networks database elaborated in ENEON

All networks and properties collected are shown and freely accessible in the website <http://www.eneon.net/graph/>.

The screenshot shows a web browser displaying the ENEON graph interface. The top part shows a complex network graph with various nodes representing different observation networks and their relationships. The bottom part shows an example of RDF encoding for a network, including metadata and relationships.

```

@context: {
  "eneon": "http://www.eneon.net/",
  "DGMARE": "http://www.europa.eu/ec/DGMARE",
  "H2020": "http://www.europa.eu/ech2020",
  "FP3": "http://www.europa.eu/ec/fp3",
  "FP5": "http://www.europa.eu/ec/fp5",
  "FP6": "http://www.europa.eu/ec/fp6",
  "FP7": "http://www.europa.eu/ec/fp7",
  "id": "@id",
  "source": "@id",
  "type": "@type",
  "network": "eneon:network",
  "http://www.eneon.net/AERONET" <http://www.eneon.net/extent> "Global" .
  "http://www.eneon.net/AERONET" <http://www.eneon.net/logo> <http://www.eneon.net/graph/logos/aeronet.jpg> .
  "http://www.eneon.net/AERONET" <http://www.eneon.net/member> <http://www.eneon.net/MAN> .
  "http://www.eneon.net/AERONET" <http://www.eneon.net/theme> "Atmosphere" .
  "http://www.eneon.net/AERONET" <http://www.eneon.net/title> "Aerosol Robotic NETwork" .
  "http://www.eneon.net/AERONET" <http://www.eneon.net/url> <http://aeronet.gsfc.nasa.gov> .
  "http://www.eneon.net/AGAGE" <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://www.eneon.net/type/Network> .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/SBA> "Public Health Surveillance" .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/description> "GAGE is distinguished by its capability to measure over the globe at high fr" .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/extent> "Global" .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/logo> <http://www.eneon.net/graph/logos/agage.png> .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/member> <http://www.eneon.net/MAN> .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/theme> "Atmosphere" .
  "http://www.eneon.net/AGAGE" <http://www.eneon.net/title> "Advanced Global Atmospheric Gases Experiment" .
  
```

ENEON graph of the EO Networks in Europe and example of the RDF encoding

2.7.3. Results in standards and metadata

The network database is maintained in a JSON file that encodes the model. It has been encoded in JSON-LD that can be dynamically converted into RDF and connected to the

linked data. The JSON syntax and semantics is validated using a JSON Schema. The JSON Schema also describes all the classes and attributes acting as a metadata resource for the JSON file.

The graph representation is based on an open source code available at GitHub called d3-process-map that uses the d3.js graphical library. The need for a PHP server side has been removed and all the “intelligence” has been moved to the JavaScript that is execute on the client side and generates and controls the representation. A relational role representation has been added as well.

2.7.4. Results in data sharing

All data collected is available through the ENEON graph website and can be downloaded in a JSON file: <http://www.eneon.net/graph>.

This is a fragment example of the object that describes the LTER network
 For instance:

```
{
  "id": "eneon:LTER-Europe",
  "type": "NetworkOfNetworks",
  "theme": [
    "Ecosystems",
    "Biosphere"
  ],
  "SBA": [
    "Biodiversity and Ecosystem Sustainability"
  ],
  "extent": "European",
  "title": "Long-Term Ecosystem Research in Europe",
  "description": "LTER-Europe was launched in 2003 as the umbrella network for Long-Term Ecosystem Research (LTER) in Europe. It's members are national networks operating a wide range of research and monitoring sites as well as larger platforms for socio-ecological research. Through research and monitoring, LTER seeks to improve our knowledge of the structure and functions of ecosystems and their long-term response to environmental, societal and economic drivers",
  "url": "http://www.LTER-europe.net/",
  "logo": "logos/LTER.jpg",
  "contact": "michael.mirtl@umweltbundesamt.at",
  "alternativeContact": "Herbert.Haubold@umweltbundesamt.at",
  "ESFRI": true,
  "relatedProject": [
    "COOPEUS"
  ],
  "links": {
    "source": "eneon:LTER-Europe",
    "nationalNetwork": [
      "eneon:LTER-Austria",
      "eneon:LTER-Belgium",
      "eneon:LTER-Bulgaria",
      "eneon:LTER-Czech",
      "eneon:LTER-Denmark",
```

```

"eneon:LTER-Finland",
"eneon:LTER-France",
"eneon:LTER-Germany",
"eneon:LTER-Hungary",
"eneon:LTER-Israel",
"eneon:LTER-Italy",
"eneon:LTER-Latvia",
"eneon:LTER-Lithuania",
"eneon:LTER-Netherlands",
"eneon:LTER-Norway",
"eneon:LTER-Poland",
"eneon:LTER-Portugal",
"eneon:LTER-Romania",
"eneon:LTER-Serbia",
"eneon:LTER-Slovakia",
"eneon:LTER-Slovenia",
"eneon:LTER-Spain",
"eneon:LTER-Sweden",
"eneon:LTER-Switzerland",
"eneon:LTER-UK"
]
},
"supportingProject": [
{
  "id": "H2020:654359",
  "name": "eLTER H2020",
  "title": "European Long-Term Ecosystem and socio-ecological Research
Infrastructure",
  "extent": "European",
  "program": "H2020-INFRAIA-2014-2015",
  "active": true
}
],
"connectedToDAB": null

```

The JSON file is automatically converted to an RDF representation (using a JSON-LD github lib) that is also downloadable.

2.7.5. Results in archiving and preservation (including storage and backup)

All data collected is kept by CREAM in the ENEON website. CREAM assumes the responsibility to store and preserve it.

2.8. ENEON's graph user feedback

2.8.1. Data set reference and name

ENEON's graph user feedback.

2.8.2. Data set description

A publicly accessible feedback catalogue integrated into the ENEON's graph (www.eneon.net/graph) coming from an extension of the GeoViQua feedback model has been set for providing feedback on missing networks or missing data on current existing networks in the graph.

2.8.3. Results in standards and metadata

An extension of the GeoViQua feedback model for geospatial data with the possibility to provide feedback for missing networks or data has been applied. The stored metadata follows the Geospatial User Feedback standard.

2.8.4. Results in data sharing

User feedback data has been made available using a REST API featuring no personal data.

2.8.5. Results in archiving and preservation (including storage and backup)

The ENEON's graph user feedback is a constantly evolving database that can be updated from the web. The feedback is stored in a S&T corp. server that originally provided the development for the GeoViQua project. No additional efforts for preserving the current status of this dataset has been made in ConnectinGEO because all relevant information has been incorporated to the ENEON graph as it arrived.

2.9. Observation gaps table

2.9.1. Data set reference and name

Observation gaps table.

2.9.2. Data set description

To better collect gaps in Earth observation data, an online editable table was prepared and made available for scientists, stakeholders and general users to populate: <http://twiki.connectingeo.net/foswiki/bin/view/ConnectinGEOIntranet/GapAnalysisTable>

EO gaps are defined by the following set of properties:

- Gap type:
 - Geographical extent (1.1)
 - Vertical extent (1.2)
 - Temporal extent (1.3)
 - Spatial resolution (2.1)
 - Vertical resolution (2.2)
 - Temporal resolution (2.3)
 - Uncertainty (3.1)
 - Geographical inconsistency (4.1)
 - Temporal inconsistency (4.2)
 - Boundary conditions issue (4.3)
 - No catalogue (5.1)
 - Catalogue saturation (5.2)

- Cannot be viewed (5.3)
- No easy access (5.4)
- Non well known format (5.5)
- No processable (5.6)
- Unknown semantics (5.7)
- No fast access to big data (5.8)
- No access (6.1)
- No open access (6.2)
- No quality (6.3)
- No provenance (6.4)
- Bad metadata (6.5)
- No metadata (6.6)
- No measured (7.1)
- No parameter (7.2)
- No interdisciplinary coordination (8.1)
- No coordination of obs. sites (8.2)
- Essential Variables
- Threads:
 - Observation requirement (1)
 - Research programs targets (2)
 - Consultation process (3)
 - GEOSS DAB analysis (4)
 - Industry-driven challenges (5)
- RS/In-Situ:
 - RS (1)
 - In-Situ (2)
 - Both (3)
- Status:
 - Detected (1)
 - Reviewed (2)
 - Accepted (3)
 - Discarded (4)
 - Prioritized (5)
 - Solved (6)
- Themes or SBAs (old ones):
 - Climate (CL)
 - Oceans (OC)
 - Water (WA)
 - Biodiversity (BI)
 - Disaster resilience (DI)
 - Energy (EN)
 - Agriculture (AG)
 - Health (HE)
 - Human settlements (HU)
 - Multiple (10)
 - All (11)
- Feasibility:
 - Very high (4): There is a mature technique
 - High (3): There was already research and maturing the technique is need
 - Medium (2): There is an idea to fill the gap that needs research

- Low (1): There is not technology foreseen to fill the gap
- Impact:
 - Very high (4): Most of the communities or topics will be impacted
 - High (3): More than one community or topic will be impacted
 - Medium (2): A community or topic is identified
 - Low (1): Not able to identify a community of topic
- Cost:
 - Very high (4): more than 20MEUR
 - High (3): more than 5MEUR less than 20MEUR
 - Medium (2): more than 0.5MEUR less than 5MEUR
 - Low (1): less than 0.5MEUR
- Timeframe to implement a solution:
 - Long term (3): more than 10 years
 - Mid-term (2): less than 5 and more than 2 years
 - Short term (1): less than 2 years
- Priority:
 - Crucial (1)
 - Very high (2)
 - High (3)
 - Medium (4)
 - Low (5)
 - None (6)

This table was created and is maintained in a collaborative twiki environment. The results of this table have been exported for further processing while the original table is still in the twiki server.

2.9.3.Results in standards and metadata

The data collected is been stored in a server. Thus, no specific standards are used. However, the export of the data for further analysis was performed using a CSV (comma separated values) structure.

2.9.4.Results in data sharing

The analysis of the gaps collected in the table was published in a statistical form as part of the ConnectinGEO *D6.2 Gap analysis draft report*. The gaps themselves are publicly available at the twiki page: <http://twiki.connectingeo.net/foswiki/bin/view/ConnectinGEOIntranet/GapAnalysisTable>

2.9.5.Results in archiving and preservation (including storage and backup)

The deliverable (D6.2) has been stored on the web site and the project digital repository. The gaps themselves are stored by CREAM in the server hosting the twiki page. The twiki platform has a versions' maintenance system allowing data preservation. CREAM as well regularly executes backup procedures.

2.10. Observation gaps user feedback

2.10.1. Data set reference and name

Observation gaps user feedback.

2.10.2. Data set description

A publicly accessible feedback catalogue integrated into the observation inventory coming from an extension of the GeoViQua feedback model has been set for geospatial data providing feedback for missing data, i.e. without a reference to a dataset.

2.10.3. Results in standards and metadata

An extension of the GeoViQua feedback model for geospatial data with the possibility to provide feedback for missing data, i.e. without a reference to a dataset has been used.

2.10.4. Results in data sharing

User feedback data has been made available using a REST API featuring no personal data.

2.10.5. Results in archiving and preservation (including storage and backup)

The Observation inventory user feedback is a constantly evolving database that can be updated from the web. The feedback is stored in a S&T corp. server that originally provided the development for the GeoViQua project. No additional efforts for preserving the current status of this dataset has been made in ConnectinGEO because all relevant information has been incorporated to the gaps table as it arrived.

2.11. Information regarding the Essential Variables

2.11.1. Data set reference and name

Information regarding the Essential Variables.

2.11.2. Data set description

The dataset contains the thesauri of EV and the indicators used in retrieving them.

2.11.3. Results in standards and metadata

No standards from vocabularies and metadata has been applied yet. The project recognizes the value of having an OWL vocabulary for EVs but the activity has been considered out of scope of the project.

2.11.4. Results in data sharing

A list of final agreed EV are provided as well as the indicators to retrieve them in the following PDF reports available in the project digital repository: *D2.2 EVs current status in different communities and way to move forward*, and *D2.3 Proposal of EVs for selected themes*.

2.11.5. Results in archiving and preservation (including storage and backup)

Report are made accessible on the ConnectinGEO website and the project digital repository, and CREAM and the UAB assume the responsibility to preserve them.

3. Report on the participation in the Pilot on Open Research Data in Horizon 2020

The Open Research Data Pilot aims to make the research data generated by selected Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access. Resulting data and products should be deposited in a research data repository where they will be findable and accessible for others.

Being part of the Pilot on Open Research Data in Horizon 2020, a data repository to preserve your data, metadata and possibly tools in the long term had to be selected.

In the case of ConnectinGEO, the data repository used has been the Digital Data Repository from the Autonomous University of Barcelona, DDD: <http://ddd.uab.cat/collection/connectingeo?ln=en>