Copernicus App Lab: A Platform for Easy Data Access Connecting the Scientific Earth Observation Community with Mobile Developers

Manolis Koubarakis
National and Kapodistrian University of Athens
koubarak@di.uoa.gr

Hervé Caumont
Terradue Srl
herve.caumont@terradue.com

Ulrike Daniels
AZO Anwendungszentrum GmbH
Ulrike.Daniels@azo-space.com

Erwin Goor
VITO
erwin.goor@vito.be

Lara König
AZO Anwendungszentrum GmbH
Lara.Koenig@azo-space.com

Valentijn Venus
RAMANI B.V.
valentijn.venus@ujuizi.com

CCS CONCEPTS

• **Information systems → Information integration**; **Spatial-temporal systems**; **Resource Description Framework (RDF)**;

KEYWORDS

satellite data, linked data, Copernicus

ACM Reference Format:

Manolis Koubarakis, Hervé Caumont, Ulrike Daniels, Erwin Goor, Lara König, and Valentijn Venus. 2018. Copernicus App Lab: A Platform for Easy Data Access Connecting the Scientific Earth Observation Community with Mobile Developers. In *Proceedings of The 2018 Web Conference Companion (WWW'18 Companion).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3184558.3186204

1 ABSTRACT

Copernicus App Lab[1] is a two year project (November 2016 to October 2018) funded by the European Commission under the H2020 program. The consortium consists of AZO[2] (project coordinator), National and Kapodistrian University of Athens[3], Terradue[4], RAMANI[5] and VITO[6]. The main objective of Copernicus App Lab is to make Earth observation data produced by the Copernicus program of the European Union available on the Web as linked data to aid their use by mobile developers.

2 WHY COPERNICUS APP LAB IS IMPORTANT AND TIMELY

Copernicus is the European program for Earth observation. It collects data about our planet using a set of satellites (the Sentinel families) and contributing missions (existing commercial and public satellites). The first satellite (Sentinel-1A) was launched in 2014 and close to 20 satellites will be deployed by 2030. Copernicus also collects information from in-situ systems such as ground stations, which deliver data acquired by a multitude of sensors on the ground, at sea or in the air. The most recent study of the European Commission predicts that the cumulative economic value of Copernicus in the years 2008-2020 will be in the range of EUR 13.5 billion.

The Copernicus services transform the wealth of satellite and in-situ Copernicus data into value-added products by processing and analysing the data. There are six Copernicus services covering the following thematic areas: Atmosphere, Marine, Land, Climate, Emergency and Security.

Data produced by the Copernicus services can be utilized in many applications with financial and environmental impact in areas such as emergency management, climate change, agriculture and security. This potential has not been fully realized up to now, because Copernicus services data "is hidden" in various archives operated by various European entities (e.g., the Flemish research institute VITO which participates in Copernicus App Lab). Therefore, a user that would like to develop an application needs to search in these archives, discover the needed data and integrate it in his application. Copernicus App Lab will demonstrate how to "break these silos open" by publishing their data as RDF, interlink it with other relevant data, and make it freely available on the Web to enable the easy development of geospatial applications.

3 SOFTWARE COMPONENTS

Copernicus App Lab will extend existing data cataloging functionality with Semantic Web standards providing a framework for explicitly describing the concepts implicit in Earth observation data and other gridded data. A core component is the Streaming Data Library (SDL), which interoperates with the Data Access Protocol in order to deliver streams out of loosely coupled Copernicus service providers. By equipping Copernicus service providers of land, atmosphere and marine data with the same distributed data access technology, streaming access can be expanded, adding to an already rich collection of gridded data offerings to application developers from various national projects, EC initiatives, and Sentinel-1/2 data as provided by ESA. The SDL is developed by partner RAMANI

[1]http://www.app-lab.eu/
[2]http://www.anwendungszentrum.de/
[3]http://kr.di.uoa.gr/
[4]https://www.terradue.com/portal/
[5]https://ramani.ujuizi.com/
[6]https://remotesensing.vito.be/

WWW'18 Companion, April 23–27, 2018, Lyon, France

ACM ISBN 978-1-4503-5640-4/18/04.
https://doi.org/10.1145/3184558.3186204

and it is currently available for a broad range of client platforms (Android, iOS, SAP-HANA XS, JavaScript etc.).

The following linked geospatial data tools developed by the University of Athens [1] are also utilized in the project to make Copernicus services data available as linked data and interlink it with other data sources:

- **GeoTriples**: an open source tool that supports the automatic transformation of geospatial data from various formats into RDF using Semantic Web standards.
- **Silk**: the well-known framework for interlinking heterogeneous data sources. Silk has been extended to support the discovery of geospatial and temporal relationships among RDF resources.
- **Strabon**: a state-of-the-art spatiotemporal RDF store that supports storing spatiotemporal RDF data and evaluating spatiotemporal queries on it, using the query languages stSPARQL and the OGC standard GeoSPARQL.
- **Ontop-spatial**: a geospatial Ontology-Based Data Access system able to integrate geospatial data sources using ontologies and mappings, and execute GeoSPARQL queries on them.
- **Sextant**: a Web-GIS tool for browsing and visualizing geospatial data available as KML, GML, image files and through WMS services, and communicating with various SPARQL endpoints to produce map layers out of the results of GeoSPARQL or stSPARQL queries.

All the above tools are open source and can be found at the following URL: http://kr.di.uoa.gr

4 PRESENTATION AND DEMONSTRATION

Our presentation will introduce the architecture of the Copernicus App Lab as managed on the Terradue Cloud Platform, and demonstrate the main technologies in a simple application. We will use a scenario from an environmental application whose aim is to study the evolution of green areas in Paris through time. We will showcase how this is achieved by using Earth Observation data in combination with other open geospatial data. More specifically, we will use the following datasets:

- The Leaf Area Index (LAI) dataset which is provided by VITO and is accessed through the OPeNDAP RAMANI server. This dataset is available as linked geospatial data through a GeoSPARQL endpoint. The LAI dataset provides observations that indicate how "green" an area is at a given time point.
- The most recent CORINE Land Cover (CLC) dataset which is available by the European Environment Agency (EEA) and contains information about the land cover of areas in Europe. Both the LAI and CLC datasets are offered by the Land Service of Copernicus.
- The Global Administrative Divisions dataset (GADM). This is an open dataset that contains information about the administrative divisions of countries worldwide. In this scenario we use the administrative divisions of Paris, which we have made available as linked data.
- The OpenStreetMap (OSM) dataset that provides information about points of interest. We will use the OSM data about France in the form of shapefiles imported in a PostGIS database. This data is available as virtual RDF graphs using the OBDA system Ontop-spatial.

All datasets are available via GeoSPARQL endpoints. Strabon and Ontop-spatial are used as the GeoSPARQL query engines, while Sextant will be used for the visualization of geospatial results on the map.

The workflow of the demonstration scenario is the following:

(1) First, the user retrieves the most up-to-date LAI data using a GeoSPARQL query. This data is accessed using RAMANI OPeNDAP API, and it gets pre-processed and converted into virtual RDF triples on-the-fly using Ontop-spatial. The results are visualized in Sextant using a different colour for each LAI range. Using the timeline feature of Sextant, we can see the areas whose LAI changes through time (the colour of the points indicating these areas changes).
(2) Then, another layer on the map will be created using the results of a GeoSPARQL query that retrieves Corine Land Cover data about Paris that is available as linked data. This dataset can be compared with the LAI dataset, as we can see the land cover of areas that appear to be green according to the LAI dataset and especially the land cover of dataset that appear to be less or more green through time.
(3) Next, we will do analytics combining the LAI, the CLC and the GADM dataset. For example, we can retrieve the average LAI value per administrative unit so that we can find out which is the"greener" administrative unit in Paris. Also, we can retrieve the most frequent CORINE land cover classes by aggregating the geometries of the areas in Paris that belong to the same CORINE land cover category and use different colour per class to visualize this on the map.
(4) Last, the OSM dataset will be employed. We will then correlate the results of the CLC and LAI datasets about green areas with points of interest from OSM that are typically located in green areas (e.g., parks, forests, etc.).

The software components we presented above have already undergone a first round of testing in the App Camp that took place on September 11-18, 2017 in Frascati at the premises of the European Space Agency[7], and more testing and experimentation by mobile developers is underway.

ACKNOWLEDGEMENT

This work has been funded by the H2020 project Copernicus App Lab, Grant Agreement number 730124.

REFERENCES

[1] M. Koubarakis, K. Bereta, G. Papadakis, D. Savva, and G. Stamoulis. 2017. Big, Linked Geospatial Data and Its Applications in Earth Observation. *IEEE Internet Computing* July/August (2017), 87–91.

[7] http://www.app-camp.eu/frascati/