# Linked Open Data in the Earth Observation Domain: the Vision of Project LEO

Manolis Koubarakis[1], Charalampos Nikolaou[1], George Garbis[1], Konstantina Bereta[1], Panayiotis Smeros[1], Stella Gianakopoulou[1], Kallirroi Dogani[1], Maria Karpathiotaki[1], Ioannis Vlachopoulos[1], Dimitrianos Savva[1], Kostis Kyzirakos[2], Stefan Manegold[2], Bernard Valentin[3], Nicolas James[3], Heike Bach[4], Fabian Niggemann[4], Philipp Klug[4], Wolfgang Angermair[5], and Stefan Burgstaller[5]

[1] National and Kapodistrian University of Athens, Greece
{koubarak, charnik, ggarbis, konstantina.bereta, psmeros, sgian, kallirroi, mkarpat, johnvl, dimis}@di.uoa.gr
[2] Centrum Wiskunde & Informatica, Netherlands
{kostis.kyzirakos, stefan.manegold}@cwi.nl
[3] Space Applications Services, Belgium
{bernard.valentin, nicolas.james}@spaceapplications.com
[4] VISTA Geowissenschaftliche Fernerkundung, Germany
{bach, niggemann, klug}@vista-geo.de
[5] PC-Agrar Informations und Beratungsdienst, Germany
{Angermair, Burgstaller}@eurosoft.de

Lots of Earth Observation (EO) data has become available at no charge in Europe and the US recently and there is a strong push for *more open EO data*. *Linked data* is a new data paradigm which studies how one can make RDF data available on the Web, and interconnect it with other data with the aim of increasing its value. In the last few years, linked *geospatial* data has received attention as researchers and practitioners have started tapping the wealth of geospatial information available on the Web. As a result, the *linked open data (LOD) cloud* has been rapidly populated with geospatial data some of it describing EO products (e.g., CORINE Land Cover and Urban Atlas published by project TELEIOS). The abundance of this data can prove useful to the new satellite missions (e.g., Sentinels) as a means to increase the usability of the millions of images and EO products that are expected to be produced by these missions.

However, open EO data that are currently made available by space agencies such as ESA and NASA are *not* following the linked data paradigm. Therefore, from the perspective of a user, the EO data and other kinds of geospatial data necessary to satisfy his or her information need can only be found in different data silos, where each silo may contain only part of the needed data. *Opening up these silos* by publishing their contents as RDF and interlinking them with semantic connections will allow the development of data analytics applications with great environmental and financial value.

Our earlier project TELEIOS (`http://www.earthobservatory.eu/`) concentrated on developing data models, query languages, scalable query evaluation techniques, and efficient data management systems that can be used to prototype applications of linked EO data. However, developing a methodology and

related software tools that support the whole life-cycle of linked open EO data (e.g., publishing, interlinking etc.) has *not* been tackled by this project. The main objective of the European project "Linked Open Earth Observation Data for Precision Farming" (LEO) presented in this paper is to go beyond TELEIOS by designing and implementing software supporting *the life-cycle of linked open EO data* and its combination with linked geospatial data, and by developing a precision farming application that heavily utilizes such data.

The scientific and technical objectives of LEO can be briefly described as follows:

1. To specify the whole life-cycle of linked open EO data and auxiliary geospatial data (e.g., maps, meteorological data) that are typically made available by public bodies and utilized in EO applications (e.g., precision farming) and publish the developed tools as an infrastructure that can be easily used by data publishers and application developers.
2. To design and implement an extraction and transformation tool that takes as input vector or raster EO data and open geospatial data and their metadata available in some well-known format (e.g., a shapefile), transforms it into RDF and makes it available on the LOD cloud.
3. To develop concepts, techniques and tools that will allow data publishers to discover geospatial, temporal and similarity relations among open EO data and other open geospatial data and metadata. The developed linking tool will be an extension of the well-known tool Silk (`http://silk.wbsg.de/`) which currently does not support the discovery of such kind of relations.
4. To develop tools for (i) cross-platform searching over linked EO metadata using keywords expressing a user information need, time predicates and spatial predicates, and (ii) a tool for browsing and visualizing time evolving linked geospatial data and the creation, sharing, and collaborative editing of 'temporally-enriched' thematic maps which are produced by combining different sources of such data and other geospatial information available in standard OGC file formats (e.g., KML). The latter tool will be an extension of the tool Sextant (`http://sextant.di.uoa.gr/`) developed in TELEIOS which will be re-developed for mobile platforms (tablets and smartphones).
5. To demonstrate the value of the developed tools by (i) performing large-scale publication and linking of open EO data from the GMES Space Component Data Access warehouse managed by ESA, and geospatial datasets made available by other public bodies in Europe, and (ii) developing a precision farming application that shows how geo-information services based on linked open EO data, linked geospatial data and specialized algorithms can contribute to an environmentally friendly increase in the efficiency of agricultural production.

**Acknowledgements.** This work has been funded by the FP7 project LEO (611141) (`http://linkedeodata.eu/`).