

Transforming EO value added products in Linked Open Data: the role of INSPIRE

Massimo Zotti, Mauro Casaburi, Marco Fago
{zotti|casaburi|fago}@planetek.it
Planetek Italia s.r.l. – Via Massaua, 12 – 70132 Bari – Italy
Tel.+39 080 9644200 - www.planetek.it

Summary

The availability of remotely sensed data acquired by satellite, airborne or UAV platforms, increases continuously, with a parallel increase in the availability of open satellite datasets. More and more in fact the space agencies data policies go in the direction of providing open access to the data collected by satellite platforms for Earth observation (EO). Those information sources bring to the *imagery big data challenge*: a growing availability of data must correspond to a greater ability to automatically and quickly extract and share information, in the most effective way.

Transforming the value-added information extracted from satellite images in Linked Open Data (LOD) responds to this need. The analysis may be fostered by crossing data from different sources of information, accessible via the Internet, which refer to the same area of interest.

The use of standardized and common data structures and semantics allows to define common ontologies: this way the information may express its full value, without the intervention of human operators. INSPIRE, the European Directive which defines the specifications for 34 data themes (hydrography, administrative units, parcels, buildings, etc.) which all Member States have to comply, resolves the problem of the heterogeneity of formats, data structures and semantics.

The exploitation of value added information, quickly extracted from Earth observation data, and compliant with the specifications defined by the European regulations, can be by far enhanced thanks to LOD, thus overcoming the inefficiencies and delays that still today restrain the real value of such an extraordinary sources of information.

This article aims to identify some business opportunities for Earth observation companies providing the so-called *downstream services* within the European program Copernicus, related to LOD and arising from the imminent availability of the Sentinel satellite data.

The European program Copernicus and the Open Earth observation Data

Copernicus¹ is the European program for the environmental monitoring that leverages the technology of earth observation to understand how our planet and its climate are changing, which is the impact of human activities on these changes, and how they will affect our daily lives.

Copernicus represent, if well used, an important tool to increase the security of European environmental policies for sustainable development, but also to foster the development of a new economy based on the use of EO data. A recent study of the European Commission² has established that, in the next 15 years, Copernicus will generate more than 20,000 new direct jobs, compared to the current 5,000 employed in the sector, especially in the *downstream* segment, that is the sector where companies offering value added services on EO data operate. This sector largely involves SMEs, small and medium-sized companies that develop commercial applications based on EO data. Indeed, according to this study, the indirect benefit offered by the integration of these data with other information (cadastral, meteorological, traffic data etc..) will have an even greater impact on the overall employment in Europe, with more than 80,000 new jobs created by 2030.

¹ <http://www.copernicus.eu/>

² SpaceTec Partners, 2012: "Assessing the Economic Value of Copernicus' European Earth Observation and Copernicus Downstream Services Market Study" - publishable Executive Summary – Final” <http://www.copernicus.eu/pages-principales/library/study-reports/>

Specifically for the operational needs of the Copernicus programme, the European Space Agency (ESA) is developing five new missions called Sentinels³: five families of satellites, scheduled for launch by 2020, equipped with synthetic aperture radar and multi-spectral sensors, for the monitoring of land, ice, oceans and atmosphere.

A wide discussion on the Sentinel data policy took place recently, with strong pressure on one side from the companies operating in the downstream sector in favor of an Open approach, and some rigidity from commercial data providers on the other side⁴. The direction however is towards a free availability, thanks to Copernicus, of accurate, frequent and open access EO information for the European states, except for a few limitations.

With this increasing availability of EO data, the so-called *big data challenge* arises. In the geospatial field, the paradigm of big data is declined by referring to the enormous volume of data globally available, and their high update frequency through new collections.

One of the greatest challenges today is the ability to put in place fast analysis processes and extraction of information from multi-temporal series of EO data, in order to provide quick and standardized quantitative information about the impact of natural emergencies (earthquakes, floods) or about the evolution of natural or human phenomena (desertification, land use, urban sprawl, etc..). Thanks to the repetitiveness of the analysis processes, statistical indexes can be extracted (eg the soil loss) useful to support the decision making.

Value-added products from EO data and Linked Open Data

Even the data policies of commercial providers of EO data now allow the Value Added Provider (typically a company operating in the aforementioned downstream segment) to freely decide the type of license preferred for its so-called derivative products. These are new geospatial products resulting from the creation of new information, on the top of the existing product, which have different properties than those of the product they are derived from, and that do not include the original product, that is the satellite image itself. A derivative product is, for example, the vector layer that represents the presence of buildings in a given area, produced recognition and classification of objects in the satellite image.

There are several examples of value-added products that can be obtained from satellite imagery⁵, and the possibility to distribute them as *opendata* creates new business opportunities for companies who overcome the traditional supply models.

Alongside of the high frequency of data provision by the current and future satellite missions, and the corresponding automation of processing of these data, the analysis of this information is nevertheless still slow. This slowness jeopardize the duration of the knowledge building process. When several information are combined and analyzed in the traditional way, with a manual analysis made by human operators, the lengthening of the access to knowledge is inevitable.

On the contrary, the deployment of automatic tools to transform these data in Linked Open Data⁶ (LOD) means that both access to the information and its exploitation becomes an automatic process, not affected by the human intervention. The publication of data as LOD is based on standard and open web technologies so that the information that can be read and understood by computers, in order to automatically connect and use data coming from different sources.

³ http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview4

⁴ EARS, 2013: "Open data study – Final", <http://easrc.org/news/easrc-study-on-the-economic-benefits-of-a-free-and-open-data-policy-for-sentinel-satellite-data>

⁵ Planetek Italia set-up a large family of value added EO products distributed with a CC-BY 4.0 license which allows their free use as open data, http://www.planetek.it/prodotti/tutti_i_prodotti/preciso

⁶ http://en.wikipedia.org/wiki/Linked_data

Along came INSPIRE

INSPIRE⁷ is the European Directive (2007/2/EC) that aims to realize the interoperability of spatial data infrastructures of the Member States and to support policies that have a direct or indirect impact on the environment. It's important to emphasize that this important project defines the rules for the sharing of data, metadata and services. These rules are well defined in the Implementing Rules, that are the regulations issued by the European Commission. These are binding legal acts, and therefore mandatory and applicable in all Member States.

INSPIRE provides the specification for 34 data themes, identified in Annexes I, II and III of the Directive. Today, therefore, those who work in the field of geomatics are supposed to know that geographic information pertaining to the INSPIRE data themes, although obtained through heterogeneous processes, must be compliant to data specifications standardized and shared at European level.

Defining ontologies for the INSPIRE data models has thus become a necessity for the interlinking of European geospatial data. It's important to highlight that INSPIRE, from this point of view, represents a unique opportunity to overcome well known problems in other areas of the Semantic Web, deriving from the extreme variety of formats, data structures and semantics. The European Directive instead already offers a well defined and common platform based on technical and legal rules.

There are already a number of actions both at European level⁸ and at national level in Italy⁹, aiming to define a common vocabulary for ontologies that refers to the INSPIRE data models. It will be equally important to enhance the experiences already made in Italy, such as the one made by the Emilia-Romagna region¹⁰, since it is consistent with the scope of generating LOD starting from geographic databases that can be updated through the use of Earth observation data.

Conclusions

The publication in form of Linked Open Data of value-added information extracted from Earth Observation data can trigger a progressing process that fully exploits the opportunities arising from the wider availability of EO data. The advantages are many and relate to a better access for citizens and institutions no longer to simple geographical information but to a real knowledge of the territory and its dynamics, to support the EU, national and local policies.

The convergence of Copernicus with the processes triggered by the INSPIRE Directive creates exciting opportunities for companies in the Earth observation downstream sector that, in various ways, are positioned along the value chain of EO data. These companies, that are mainly SMEs, may offer innovative services in the different stages of the process: data access, processing and sharing. They can for example facilitate the intelligent access to data in emergency situations or to catalogs for the best selection of the scenes of interest; processing the images for the extraction of value-added information, through rapid standardized procedures that ensure standardized products comparable over time; promote the sharing in interoperable formats and structures to comply to the European regulations; transform value added products in LOD, so that they can be linked to other data and further enriched, thus making room for the creation of new value-added applications, to the benefit of SMEs and the resulting creation of new jobs in Europe for high level professionals.

⁷ <http://inspire.ec.europa.eu/>

⁸ A. Perego, 2012: "Cross-domain interoperability for EU spatial data", http://www.w3.org/2012/06/pmod/pmod2012-jrc-andrea_perego.pdf; M. Lutz, A. Perego, M. Craglia, 2013: "Interoperability of (open) geospatial data: INSPIRE and beyond", http://www.w3.org/2013/04/odw/odw13_submission_58.pdf

⁹ <http://www.digitpa.gov.it/notizie/linee-guida-open-data-interoperabili>

¹⁰ Association OpenGeoData, 2013: "Dati geografici aperti: istruzioni per l'uso"

http://www.opengeodata.it/index.php?option=com_jdownloads&Itemid=0&view=summary&cid=4&catid=3