

The MELODIES project: Exploiting Linked Open Geospatial Data

*Jon Blower, University of Reading, UK
(on behalf of the MELODIES consortium)*

The European Open Data Strategy establishes important new principles that ensure that European public sector data will be released at no cost (or marginal cost), in machine-readable, commonly-understood formats, and with liberal licences enabling wide reuse. These data encompass both scientific data about the environment (from Earth Observation and other fields) and other public sector information, including diverse topics such as demographics, health and crime. Many open geospatial datasets (e.g. land use) are already available through the INSPIRE directive and made available through infrastructures such as the Global Earth Observation System of Systems (GEOSS). The intention of the Open Data Strategy is to stimulate the growth of research and value-adding services that build upon these data streams; however, the potential value inherent in open data, and the benefits that can be gained by combining previously-disparate sources of information are only just starting to become understood. **The MELODIES project (Maximising the Exploitation of Linked Open Data In Enterprise and Science) is developing eight innovative and sustainable services, based upon Open Data, for users in research, government, industry and the general public in a broad range of societal and environmental benefit areas.**

MELODIES (<http://melodiesproject.eu>) is a European FP7 project that is coordinated by the University of Reading and has sixteen partners (including nine SMEs) from eight European countries. It started in November 2013 and will run for three years. The project is therefore in its early stages and therefore we will value the opportunity that this workshop affords to present our plans and interact with the wider Linked Geospatial Data community.

The project will develop eight new services¹ covering a range of domains including agriculture, urban ecosystems, land use management, marine information, desertification, crisis management and hydrology. These services will combine Earth Observation data with other open data sources to produce new information for the benefit of scientists, industry, government decision-makers, public service providers and citizens. The long-term sustainability of the services will be assessed critically throughout the project from a number of angles (technical, political and economic), in order to ensure that the full benefits of the MELODIES project are realised in the long term. The priority of the project, therefore, is to demonstrate that releasing data openly leads to concrete commercial and scientific benefits, and can stimulate the production of new applications and viable services.

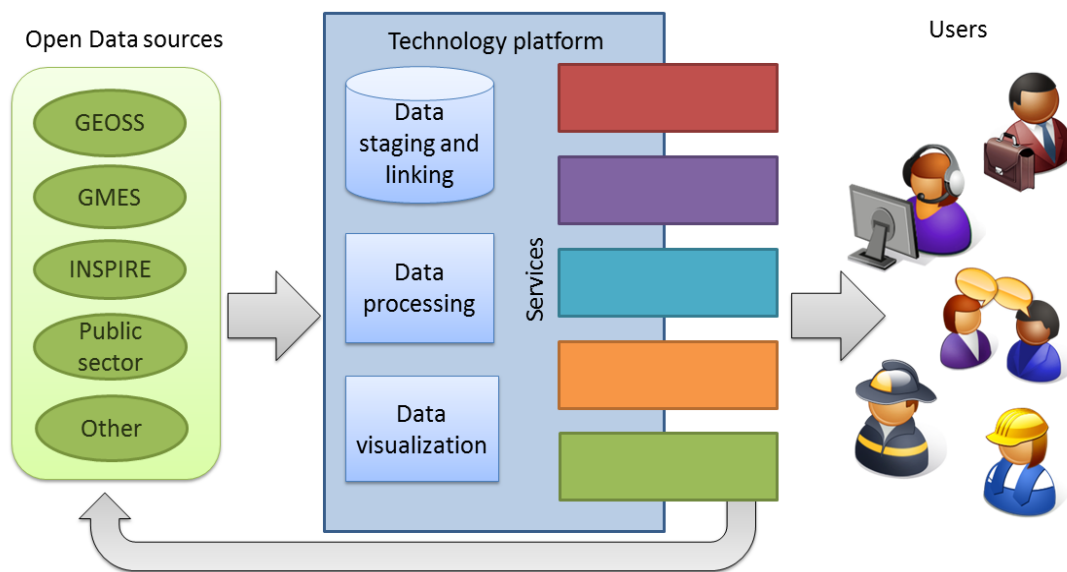
The services will be underpinned by a shared technology platform, which provides capabilities for cloud computing (using technologies developed within the GeoWOW² project), Linked Data management (using the Strabon³ data store from the University of Athens) and mapping/visualization (using various free and open-source tools). The emphasis of the project is on the re-use of existing high-quality technologies in order to enable the service developers to create robust, scalable and flexible services based upon large quantities of diverse open data.

¹ <http://www.melodiesproject.eu/services.html>

² <http://www.geowow.eu/>

³ <http://www.strabon.di.uoa.gr/>

The figure below presents a graphical overview of the project. In addition to consuming open data, the MELODIES services will also produce new open data sources, which will be fed back into existing data infrastructures.



Although the eight MELODIES services are highly diverse (responding to different kinds of user needs in different sectors), Linked Data is a common theme throughout the project. We see Linked Data as a means of breaking out of existing data silos and providing technologies to integrate data from very different sources (e.g. from Earth Observation missions and from socioeconomic databases). However, the area of Linked Data is very new to most of the project participants; therefore we anticipate facing a number of key challenges and questions:

1. Is Linked Data primarily a method for publication and data discovery, or can it really be used efficiently in the integration and analysis of large volumes of actual data?
2. How can we use Linked Data to record, exploit and publish the provenance of datasets?
3. How can Linked Data services interoperate with other open standards, such as Open Geospatial Consortium services?
4. What is the most efficient way to handle large quantities of raster data such as Earth Observation imagery in a Linked Data framework?
5. Can we use Linked Data to drive interactive and flexible data visualizations?
6. How can *time-varying* geospatial data be handled in a Linked Data environment?
 - (We plan to use stRDF and stSPARQL, which add temporal capabilities to other standards such as GeoSPARQL. What are the implications of this?)
7. Can we use Linked Data to provide a means for the user to annotate geospatial datasets, enabling the sharing of discoveries and expertise?

Although we have ideas and plans in these areas (derived from our experience with previous projects, notably TELEIOS⁴ and CHARMe⁵), we would very much welcome the opportunity to discuss these plans with the wider community, to gain early feedback and to put forward some new use cases that are perhaps relatively rarely addressed in Linked Geospatial Data projects.

⁴ <http://www.earthobservatory.eu/>

⁵ <http://www.charme.org.uk/>