

# Open Data Monitor Project

Yunjia Li, Elena Simperl, Qiong Bu, and Seyi Feyisetan

University of Southampton  
<http://www.southampton.ac.uk>

The OpenDataMonitor (ODM) is a two-year STREP project that started in November 2013. It analyses and visualises open data catalogues across Europe at the European, national, and local levels. The increasing recognition of the strategic importance and benefits of open data has been driving several efforts and initiatives to facilitate the data publishing and consumption. Open data has contributed to improving governmental transparency, and fostering future businesses. Numerous public and private sectors in different levels (European, national, local, etc) have already published or plan to publish open data in different domains.

The European Union is at the forefront of the open data innovation, and millions of data assets have been made available from EU members. However, on the pan-European level, there is a lack of analysis of open data to identify gaps and develop sustainable strategies to empower the reuse of open data. It is still difficult to overview the current status of open data in the EU and answer questions such as (1) which catalogues are available in a certain country and (2) what policy fields are covered by a particular catalogue. Hence, there is a need for a solution that can help various stakeholders involved in the open data publishing consumption process to have an up-to-date, consolidated and comprehensive view of the existing data catalogues and datasets. More importantly, we need a platform to monitor evolution of each data catalogue over time automatically, and to search and compare available resources across borders and sectors.

To solve the problems described above, ODM has developed sophisticated methods to scan data catalogues implemented in both CKAN and other platforms, analyse meta-data and provide comprehensive visualisations to compare existing open data resources. The main research and development activities in ODM project can be described as follows:

- (1) Fundamental research studies on the open data landscapes from both technical and non-technical point of views, which underpins the ODM platform design and implementation.
- (2) Elaborate a scalable open data harvesting and monitoring implementation by collecting metadata from different open data catalogues. This includes the design and realisation of a stable architecture for metadata harvesting and integration.
- (3) Develop a flexible open data analysing and monitoring tool with intuitive

dashboards and visualisations.

In Feb 2015, the alpha version of ODM platform has been released. In this EU networking session, we will demonstrate the dashboards and advanced benchmarking features. We have designed the dashboard to visualise open data metadata on pan-Europe and national levels. In the European dashboard, an aggregated view of open data analysis across EU is available based on selected metrics, such as open license, machine-readability and metadata completeness. In the national level dashboard, we provide a detailed analysis of top license types and top data distribution formats, which have been adopted in data catalogues. For some selected countries, 4 quantitative and 5 qualitative metrics will be available to show the rankings of the selected countries. We will also have an advanced benchmarking feature to allow users to customise the data visualizations and generate report based on selected countries, catalogues and metrics.

The EU Project Networking Track provides a unique opportunity for ODM to engage with other projects. From the last one and half years of R&D activities in ODM, we have gained a lot of experience on the landscape of open data in EU and what data visualization methods could be used to reflect more insights into the big picture of open data in different countries. We will disseminate our R&D results to a larger group of audiences in this networking track. In detail, we can share our knowledge and contribute to similar projects from the following aspects:

- (1) We have manually compiled a list of open data catalogues for each member country in the EU, including their URLs, number of datasets and the policy fields that each dataset covers. We have made the list publicly available for everyone to check and update, which may help the projects that need such a list.
- (2) ODM has implemented a metadata harvesting and integration platform. In this process, we have gained a lot of insight into the diversity of the current status of open data. For example, open data catalogues are implemented using various software platforms and metadata standards. The quality of the metadata is also different from catalogue to catalogue. Our experiences in this area can advise the projects working on similar subjects.
- (3) We have documented a thorough list of metrics that could be used to analyse the status of an open data catalogue. Some of the metrics have been implemented in the platform, which is a reference to other projects for intuitive data visualisation.

ODM is looking for partnership with other projects to help shape the future of open data research. In particular, we are expecting to collect feedbacks on the current open data monitoring implementation and visualisations, so that we can improve them before the end of the project. We are also looking for potential projects to use ODM platform for open data research and may take the project forward to another stage, which is very important for the sustainability of the platform after the project is finished.