OpenDataMonitor

Publishable summary (Project period 2)

The OpenDataMonitor (ODM) project offers a new concept that provides users with a platform, which gives an overview of open data resources and undertakes both an analysis and visualisation of existing data catalogues. The tool offers a comprehensive open data analysis for open data publishers, consumers and further stakeholders all across Europe. The consortium partners deliver highly innovative methods to scan data catalogues, analyse meta-data and create comprehensive visualisations from a vast pool of resources, ranging from open data collectors on the national level to pan-European open data hubs. By using standardised APIs (e.g. CKAN) and complex harvesting mechanisms, it is possible to statistically analyse and present various metrics including file formats, machine readability, licenses and additional meta-data. Such findings can be exploited to identify the status, gaps and potentials of open data resources. Furthermore, the OpenDataMonitor platform features intuitive dashboards to enable comprehensive open data analyses for open data publishers and key stakeholders across Europe.

To achieve the stated aims, the project framework was structured into four stages. Initially the partners conducted extensive **research studies and stakeholder analysis**. This stage included a detailed investigation of current open data topologies, catalogues, metadata harmonisation, innovative monitoring methods, architectures and standards, best practice data visualisations, dashboards and key figures. During the second stage (**concept design and software development)**, the technical specifications and functionalities as well as the detailed interface designs and dashboards for the metadata harmonization and monitoring were defined. During the next phase, a demonstrator of the platform was deployed to perform usability tests and to undertake feedback loops with selected end users. The last stage of the project contained the **dissemination and exploitation** activities. Partners of the consortium disseminated the outcomes of ODM to community groups and users through a range of tools and networks such as websites, project branding, publications, presentations and public events.

A ODM RESEARCH

Throughout the second year of the ODM project, the consortium has focused on getting a clearer understanding of the experiences, challenges and requirements of open data publishers and users. The research enabled us to identify patterns of understanding and interests in order to serve these stakeholders' information needs through the OpenDataMonitor platform. The research showed that the various stakeholders' views are relatively aligned. Stakeholders across the board – whether within or outside public administration or whether they are active at the policy-level or the operational level of open data – attribute similarly high importance to aspects such as e.g. open formats, open licenses, machine-readability and that data are made available free of charge. That aspect of delivering data with meaningful and high quality metadata is seen as only slightly less important. This technically-oriented understanding of open data is further substantiated by the

findings of the second research strand. Both the survey and the follow-up interviews show that timeliness and provenance of the data are quite important also for commercial stakeholders. Equally important are the licensing schemes attached to the open datasets. Furthermore, the list of open data resources, platforms and APIs, which was setup in the first year, was extended throughout the second project period. The list of open data catalogues covers almost all countries in the EU, as well as catalogues/portals at different levels (pan-Europe, national and local levels). It serves as the primary resource for ODM data harvesting. The list of literature resources covering scientific papers, blogs, reports and other materials that describe open data from different perspectives has increased from 400+ to 500+ in this second iteration. The result is available for the public and can be reused by other projects or people who are interested in open data.

B ODM TECHNICAL DEVELOPMENT

Parallel to the comprehensive research activities that were performed during the first and second project year, the technical development took place. Essentially, the goal of the open data monitoring framework is to enable and support the harvesting of metadata from open data catalogues, their storage, processing, and analysis, and their use for generating visualizations. The detailed analysis of stakeholders enabled the creation of precise use cases. In order to implement the gathered user stories, detailed functionalities and indicators were identified. The metrics defined in the first project period were reviewed with real data harvested in the second period. Finally, the metrics were grouped based on two principles: quantitative and qualitative. The overall architecture and the processing workflow for the OpenDataMonitor system, which was designed in the first year of the project, were consequently improved. Furthermore, new functionalities and enhancements were implemented for the main components of the platform. Based on the stakeholder workshops it became apparent that the usability of the platform had increased compared to the alpha version. In the second year of development the main application features were implemented, including the three level dashboards (European, Country and Catalogue). The Benchmarking dashboard of catalogues and countries was designed and finished during the last project phase. The second year of design and implementation improved the usability of the platform, whereas the first year's prototype was used to lead the research and innovation for the missing values in open data. During the second project period, we verified the metrics and refined the definition and calculation accordingly in order to use the most relevant metrics for measuring the performance of selected entities in both the country and the catalogue level. The code of the OpenDataMonitor is available in a public GitHub repository: https://github.com/opendatamonitor.

The initial prototype was subsequently adjusted and improved based on the feedback received by the end-users. From the perspective of the usability study, the effectiveness, efficiency and satisfaction of the ODM platform's usability was evaluated along both non-functional and functional aspects. Function-wise, it was investigated whether the features provided in the ODM platform are complete, proper or redundant/unnecessary, determining to what extent the platform was beneficial

to the users. Non-function wise, it was investigated whether the user interface (such as the graphs or dashboards displayed) were easy to understand and whether the tool created a good user experience, by making it easy to explore and navigate. The results of this study were used to provide feedback for adjusting, correcting errors, and improving the ODM tool. In order to support the community, intuitive and comprehensive manuals and tutorials were created to explain and demonstrate the various components and features provided by the ODM platform, so that it can be further developed and used by a wider audience. The user manuals took into consideration what the potential use of each component and feature is, who the potential users are, and what level of expertise might be required. A decision was made on emphasizing the creation of entertaining short bits of information on the platform with pictures and videos, rather than long documents. For each selected component, a HTML-based manual with screenshots was produced as well as a complementary screencast tutorial. The official showcase gives an overarching view and shows the main functionalities of the project prototype. The ODM Knowledge Base was updated with the final support materials and further insights the ODM consortium generated throughout the project to ensure the sustainability of the project outcomes.

C ODM IMPACTS AND MEANS OF ACHIEVEMENT

In order to reach and engage with relevant communities, a dissemination and impact strategy was executed to connect the research outputs and created services or tools with the target audiences. Throughout the entire project, offline and online methods of engagement were utilized to create an ODM community. This includes the project website and a Knowledge Base, the design of a project logo and a project branding, press releases, news stories and blog posts, newsletters, information mailing, postcards and stickers, use of social media, presentations and workshops at international conferences, hackathons and meetups, scientific publications and the OpenDataMonitor symposium. Similar to building an interested community, it was important for us to engage with existing data communities.

OpenDataMonitor aims to overcome some of the main challenges in understanding the availability and gaps in open data. ODM delivers a tool for a wide range of different end user types with different purposes concerning open data. These users include: start-ups and entrepreneurs, policy makers, journalists, researchers and academics. Both publishers and consumers benefit from the tool, which is delivering:

- a better understanding of which datasets are available at the regional, national and European levels and to support an increased reuse of these open data resources by a wide range of users, from developers, to government and public bodies and citizens.
- a sharper overview of the availability of both regional and national open data, as well as a clear collection of open data resources.
- tools that support development of both sustainable and profitable open data policies and strategies, which will be enhanced by an understanding of the gaps in datasets and areas to focus on.

D ODM CONSORTIUM

SYNYO GmbH

SYNYO GmbH, an SME based in Vienna, Austria, is an independent social research and innovation hub that focuses on the intersection between Information & Communication Technologies (ICT) and Society to provide smart solutions to citizens, consumers and decision-makers.

Athena Research and Innovation Center

The Institute for the Management for Information Systems (IMIS) is a part of the Athena Research Center. The mission of IMIS is to conduct research, develop innovative applications, and provide services in the areas of large scale information systems, with a focus on geospatial data, Linked Data and Big Data.

University of Southampton

Electronics and Computer Science (ECS) at the University of Southampton has a leading role in developments in the areas of Web Science and Linked Data, including the prestigious Web and Internet Science research group, which will be directly involved in the project.

Open Data Institute

The Open Data Institute (ODI) was officially launched in December 2012. As an independent, non-profit, non-partisan company, the ODI has a funded commitment to work with the UK Government (via the UK innovation agency, the Technology Strategy Board) over the next five years.

IfG.cc

IfG.CC is a non-partisan, non-profit research institute for new forms of government and public administration in the information age. It researches and develops new ICT-based organisational models and forms of governance in a variety of policy fields.

Red.es

Red.es is the public corporate entity attached to the Ministry of Industry, Energy and Tourism (MINETUR) which is responsible for promoting the development of the Information Society in Spain. Red.es manages the Spanish Public Administration Data Reuse Catalogue, available on datos.gob.es.

City of Munich

The Munich city administration is responsible for social welfare, business and tourist promotion, services of general interest, environmental issues, cultural affairs and city planning. The Munich Managerial Board - ICT strategy - is responsible for internal IT solutions and for providing public access to data of general interest.

Contact details:

Email:	office@opendatamonitor.eu
Twitter:	@opendatamonitor
Project Website:	http://project.opendatamonitor.eu/
Demonstration Site:	http://opendatamonitor.eu/

OpenDataMonitor is co-funded by the European Commission under FP7 under grant number 611988