The Minerals4EU Knowledge Data Platform for managing Web-based mineral resources information in Europe

Daniel Cassard¹, François Tertre¹, Guillaume Bertrand¹, Agnès Tellez-Arenas¹, Frands Schjøth², Jørgen Tulstrup², Tjerck Heijboer², Jouni Vuollo³, Dana Cáповá⁴, Jasna Šinigoj⁵, Stephan Gruijters⁶, Robert Tomas⁷, Chris Schubert⁷

¹Bureau de Recherches Géologiques et Minières (BRGM), Orléans, France
²Geological Survey of Denmark and Greenland (GEUS), Copenhagen, Denmark
³Geological Survey of Finland (GTK), Rovaniemi, Finland
⁴Czech Geological Survey (CGS), Praha, Czech Republic
⁵Geological Survey of Slovenia (GeoZS), Ljubljana, Slovenia
⁶Geological Survey of the Netherlands (TNO), Utrecht, the Netherlands
⁷European Commission Joint Research Center (JRC), Ispra, Italy

Corresponding author: D. Cassard (d.cassard@brgm.fr)
BRGM, GeoResources Division
3, avenue Claude Guillemin – BP 36009
45060 – Orléans CEDEX 2 - FRANCE
Phone: +33 2 38 64 36 98

The EU-FP7 project Minerals4EU ‘Minerals Intelligent Network for Europe’ is developing an operational data management distributed system based on high-level interoperability standards using advances made in EU-FP7 former projects such as OneGeology-Europe, ProMine and EuroGeoSource, and ongoing projects like EURARE and InGeoCloudS, in terms of database structure, harvesting systems, web services, metadata management, integration of non-structured information and cloud computing. Minerals4EU will also contribute to implement the standards developed jointly by Member States and the European Commission in the framework of the INSPIRE Directive as well as extended domain specifics of a European geoscientific data infrastructure defined in the EU-FP7 EGDI-Scope project. The objective is to develop a Minerals Knowledge Data Platform (the EUMKDP) allowing to easily combine information related to primary and secondary mineral resources and to provide end-users with all the available information from primary sources to waste streams, from exploration to production and trade, from estimates of resource availability to foresight studies on raw materials supply and demand in the EU. The EUMKDP will thus represent one of the first bricks of the future European geological data infrastructure. The proposed technical solutions assure an effective and sustainable system designed for facilitating data updates and maintenance, and for giving a full access to information related to the whole mineral resources value chain.

Data related to raw materials, either metallic ores, or industrial minerals or construction materials, of primary origin or from mining and industrial wastes are, most of the time, available in Europe. However, they are often scattered amongst a variety of institutions, including governmental agencies, universities, NGOs and industries. These data are often stored in databases with their own design/architecture and vocabulary, making any attempt
of merging in view of a compilation difficult and time consuming. Problems regarding availability, quality, organization, accessibility and sharing of data are common to a number of policies and are experienced across the various levels of public authorities. Solving these problems requires measures that address exchange, sharing, access and use of interoperable spatial data and services both at national and European levels. This is the aim of the INSPIRE Directive (2007, ‘establishing an Infrastructure for Spatial Information in the European Community’), but its implementation in the Member States has just started and achieving those objectives represents a major challenge.

In order to create a system that is kept up-to-date, Minerals4EU adopted and developed a distributed architecture based on a central harvesting database synchronized with a central diffusion database. The first one controls the quality of the data coming from the data providers (essentially national and/or regional geological surveys), and the second one is optimized for diffusion, the synchronization between the two databases being made using SQL scripts. The use of central databases allows minimizing the drawbacks of a distributed architecture, improving the performance of the system, and reducing the risk of having actually inaccurate results if local services are down or unreachable. However, not all the knowledge is coming from structured data stored in databases. The system is thus also designed to accommodate semi- and non-structured information (e.g., syntheses and statistics in the form of text, graph charts and time-series, and related to exploration and primary reserves and resources, exploitation technologies including ore beneficiation, extraction technologies, endproduct development and waste management practices, European market survey and raw material demand…) in various formats (e.g., Excel spreadsheets, text files, PDF files, images…).

The system must be able to retrieve all pertinent information, spatial and nonspatial, which is related to a topic, becoming thus an incomparable tool for addressing complex questions or problems. This goes through (i) the use of metadata with the development of a specific instance of the MICKA catalogue (used in OneGeology-Europe) for geospatial data and the use of the DUBLIN CORE Metadata Element Set for other data, (ii) the use of a powerful indexation engine which will be able to index the data in the diffusion database, the documents present in the repository and their metadata. Semi- and non-structured information will be processed and indexed, first extracting relevant information from the documents (e.g., named-entity recognition on atomic elements such as names of locations, expressions of time, thematic specific terms…), then classifying this information according to three facets (spatial, temporal and thematic) and creating indexes foreach facet and (iii) the development of full-search facilities which will allow the user to search in the diffusion database, in some other databases (manually plugged into the indexation engine) and in documents added to the EU-MKDP using a dedicated interface. The user will be able to search (i) using full text or (ii) using a specific interface dedicated to the concept the user is looking for. Selected data will be sent to the map viewer of the EU-MKDP portal which will allow the user to crossreference them with other data (geological map, consumption area…).

Besides continental primary and secondary mineral resources, the EU-MKDP will also manage information related to continental resources at depth and offshore resources. In
In order to better cover the mineral resources domain, an extension of the platform to urban wastes (WEEE, ELV's and batteries) which represent an important source of secondary critical raw materials has been designed (H2020 ProSUM project recently accepted). From the technical point of view, the foreseen developments bear on domain ontologies which would dramatically improve the power of the search module, the pertinence and the exhaustiveness of the results obtained being the major contribution of a knowledge base. Looking toward the immediate future, the EU-MKDP (i) has been designed to become the core component of the EURMKB (European Union Raw Materials Knowledge Base) that the European Commission aims to get developed rapidly, and (ii) has thus the vocation to accommodate and exploit further datasets.

This presentation describes the activities being undertaken in one of six Work Packages, alongside the establishment of the EU Mineral Intelligence Network, the development of the most comprehensive European Minerals Yearbook and the compilation of a Foresight Study examining issues surrounding minerals supply and demand in Europe. Further information about the project can be found on the website www.minerals4eu.eu.

The Minerals4EU ‘Minerals Intelligence Network for Europe’ project was funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 608921.