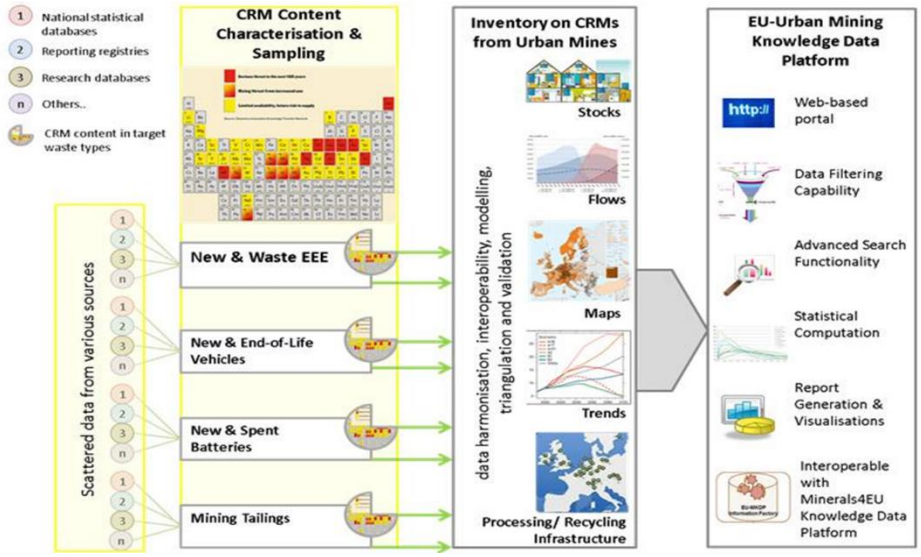
Note accompanying the delivery of the EU-UMKDP



Project	Prospecting Secondary raw materials in the Urban mine and Mining waste
Acronym	ProSUM
Grant Agreement	641999
Funding Scheme	Horizon 2020
Webpage	http://www.prosumproject.eu/
Work Package	Work Package 5
Work Package Leader	Daniel Cassard, BRGM
Deliverable Title	Note accompanying the delivery of the EU-UMKDP
Deliverable Number	D5.7
Deliverable Leader	Daniel Cassard, BRGM
Version:	1.0
Status	Final
Author	Daniel Cassard, François Tertre (BRGM), Frands Schjøth, Tjerk Heijboer (GEUS), Dana Čápová (CGS), Anders Hallberg (SGU), Jasna Šinigoj (GeoZS)
Reviewed by	Project Management Team
E-mail	d.cassard@brgm.fr
Scientific Coordinators	Jaco Huisman, Scientific Advisor, UNU – IAS SCYCLE Nikolaos Arvanitidis, Head of Unit, SGU
E-mail	Huisman@unu.edu ; Nikolaos.Arvanitidis@sgu.se

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 641999.

Document Control

Project Coordinator:	WEEE Forum
Work Package Leader:	BRGM
Deliverable leader:	BRGM
Due date:	M34 30 September 2017
Date of submission:	M34 30 September 2017
Dissemination level:	PU (Public)

Version history			
Ver. no.	**Date**	**Reason for release**	**Responsible**
V1	03/10/2017	First draft release	D. Cassard
V1	22/11/2017	Review	T. Heijboer
V1	22/11/2017	Corrections	D. Cassard
V1	14/11/2017	Description of Urban Mine Applications	F. Tertre
V1	15/12/2017	Final review	D. Cassard

Notice

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

PURPOSE

The aim of this note which accompanies the delivery of the EU – Urban Mining Knowledge Data Platform (the EU-UMKDP) is to provide the end user with a complete, but simplified and accessible overview of the whole system from 'data collection' to 'data publication' on the Web. Detailed technical approaches have been published in two public reports (Cassard et al., 2016 and Heijboer et al., 2017) accessible on demand.

This note focuses on applications running on the top of the Diffusion database and which are the most visible and useful part of the project for the end user.

The EU-UMKDP is a stand-alone platform. However it is strongly connected with some other platforms an notably the EU – Minerals Knowledge Data Platform, the EU-MKDP developed in the frame of the Minerals4EU project: The EU-MKDP provides the EU-UMKDP with information related to primary mineral resources and also mining wastes. This latter topic (data model, web service and harvesting) has been improved by ProSUM, thus showing the strong synergy between the two projects.

The EU-UMKDP, like other knowledge data platforms (KDPs), will also be connected to the MICA Expert System currently being developed in the H2020 MICA project. This connection will also be presented here as it should reinforce the impact of the information delivered by ProSUM, when it will become operational.

PROVISONAL URLs:

The EU-UMKDP: http://prosum.brgm-rec.fr/

The homepage of ProSUM applications: http://prosum.brgm-rec.fr/alphapps/prosum/homepage

EXECUTIVE SUMMARY

The H2020 project ProSUM 'Prospecting Secondary raw materials in the Urban mine and Mining waste' has developed an operational data management distributed system based on high-level interoperability standards. Using the technical and organizational advances made in former EU-FP7 projects such as OneGeology-Europe, ProMine, EuroGeoSource and Minerals4EU, and in coordination with the ongoing EURARE project, ProSUM contributes to implement the standards of a European Geoscientific Data Infrastructure defined in the EU-FP7 EGDI-Scope project. The objective was to develop an Urban Mining Knowledge Data Platform (EU-UMKDP) to provide end users (i) with all the available data/information related to wastes from the urban mine and from mining activities and (ii) with results of a comprehensive inventory identifying, quantifying and mapping CRM (Critical Raw Materials) stocks and flows at national and regional levels across Europe and to allow them to easily view and exploit these data, and combine them with primary raw materials data from the Minerals4EU project. The EU-UMKDP is thus designed as a fully-fledged extension of the EU-MKDP (EU-Minerals Knowledge Data Platform) developed within Minerals4EU and which aims to become the future European Mineral Resources (s.l.) data infrastructure. This assures an effective and sustainable system designed for facilitating data updates and maintenance and allowing full access to information related to the entire resources value chain.

The aim of this note which accompanies the release of the final version of the EU-UMKDP is not to re-describe in detail all the 'mechanics' which is behind this geospatial portal and its applications. This has already been done in two detailed public reports[1,2] published in the frame of ProSUM Work Package 5:

- *The ProSUM EU-Urban Mining Knowledge Data Platform Specification (T5.3.1) by Daniel Cassard, François Tertre, Frands Schjøth, Tjerk Heijboer, Dana Čápová, Anders Hallberg and Jasna Šinigoj (2016), that reviews all main components of the system and their connection, from data types to data models, the INSPIRE compliance, the construction of databases, web services and the harvesting system, as well as the diffusion system and the metadata catalogues for both structured and non-structured data, and the portal functionalities.*

- *Deliverable D5.5: Data models and code lists by Tjerk Heijboer, Frands Schjøth, Martin Podboj, Daniel Cassard, Francois Tertre and Anders Hallberg (2017), that details the structure of the data that will be stored in the different databases exploited by the ProSUM Urban Mine Knowledge Data Platform (EU-UMKDP) architecture.*

This note will reuse (updated) parts of these reports in order to recall what are the main technical characteristics of the EU-MKDP and will then focus on two points which have not been tackled before:

- *The applications running on the top of the ProSUM Diffusion database and which are the very reason for which the EU-UMKDP has been developed.*
- *The connection of the EU-UMKDP with other Knowledge Data Platforms (KDPs), and notably with the MICA Expert system (the EU-RMICP).*

Figure 1 below shows that Integration of all these KDPs is already on the way.They already share several components (notably the Harvesting, DB and the Diffusion DB) and they represent some of the bricks of the future European Union Raw Materials Knowledge Base (the EURMKB) that will be exploited by the future Raw Materials Information System (RMIS 2.0) currently developed by the European Commission DG JRC in Ispra.

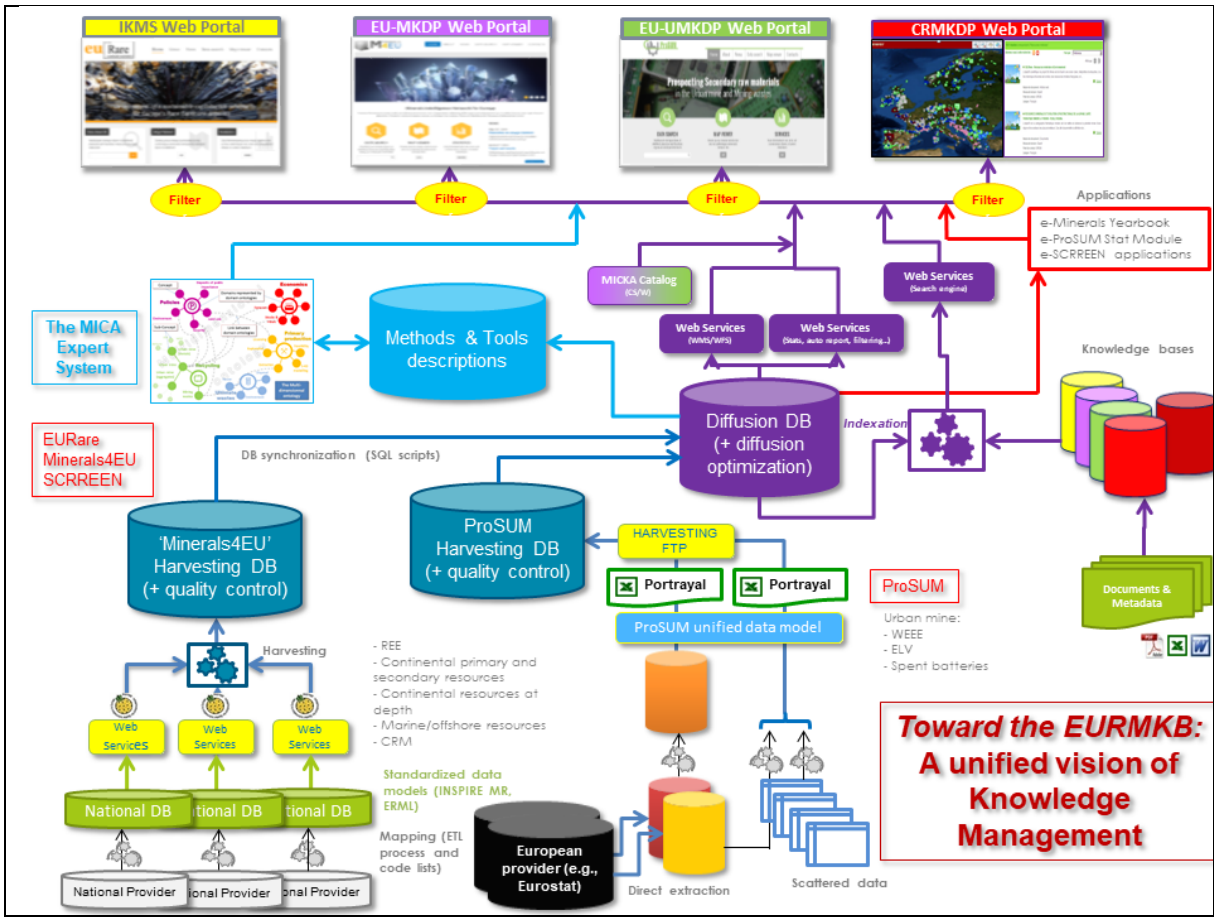


Figure 1 - The EU-UMKDP and the other KDPs. A unified vision of Knowledge Management in the Raw Materials domain, and the first steps for building of the EURMKB.

GLOSSARY OF KEY-TERMS USED

Catalogue services support the ability to publish and search collections of descriptive information (metadata) for data, services, and related information objects. Metadata in catalogues represent resource characteristics that can be queried and presented for evaluation and further processing by both humans and software. Catalogue services on the Web (**CS/W**) are required to support the discovery and binding to registered information resources within an information community. (OGC definition: http://www.opengeospatial.org/standards/cat).

Data model: A **data model** organizes data elements and standardizes how the data elements relate to one another. The ProSUM unified data model, as well as the **INSPIRE MR data model**, is an entity-relationship model (ERM). This is an abstract conceptual data model (or semantic data model) used in software engineering to represent **structured data**. This model is transformed into a relational model, which in turn is implemented in a relational database management system. These conceptual entity/relationship models are developed using **UML** Class Diagram notation. Thus, **data models describe the structure, manipulation and integrity aspects of the data stored in data management systems such as relational databases. They typically do not describe unstructured data**, such as word processing documents, e-mail messages, pictures, digital audio, and video.

Diffusion database: The role of the (Central) Diffusion Database (DB) is to provide the ProSUM portal with data. These data are sent to the portal using Web services (**WFS, JSON**). In order to speed the process, the structure of the Diffusion DB is optimized for diffusion. This means that its structure does not follow exactly the **data model** which has been 'flattened' or simplified without altering the data. For this project, the (Central) Diffusion DB is hosted by BRGM in France. At the origin – before optimization – the **Diffusion DB** is acopy of the **Harvesting DB** made by using **SQL** scripts.

Distributed architecture means that the data served by the platform are regularly uploaded from data providers (national data provider or EU provider…) through a 'harvesting' system using web services. Data are sent to a central database (DB) (actually composed of two distinct DBs, one being dedicated to harvesting, the other one to diffusion) which only acts as caching mechanisms. The central database is used to minimize the drawbacks of a pure distributed architecture: a user of the system may search for occurrences of commodities throughout Europe; in case all information is available on distributed servers, such a query will have to be executed at every data provider, resulting in a high risk of low performance. Therefore the data is stored centralized to act **as an optimized search index**. It also reduces the risk of having actually inaccurate results if local services are down or unreachable.

ERML or **EarthResourceML** data model: This is the 'international – world-wide' fully compliant version of the **INSPIRE MR** data model (http://www.earthresourceml.org/). This data model is managed by the IUGS/CGI/ERMLWG and used in Europe, North America and Australia.

ETL: Extract, Transform and Load process in database management that performs data **extraction** from homogeneous or heterogeneous data sources; data **transformation** for storing in the proper format or structure for the purpose of querying and analysis; and data **loading** into the final target.

EU-MKDP: the European Union Minerals Knowledge Data Platform developed in the frame of the EU-FP7 Minerals4EU project.

EU-UMKDP: the European Union Urban Mine Knowledge Data Platform developed in the frame of the H2020 ProSUM project. This report defines the technical specifications of this platform. This platform communicates with the EU-MKDP through web services (mostly **WFS** & **WMS**).

EU-RMICP: The European Union Raw Materials Information Capacity Platform developed in the of the H2020 MICA project. This Platform **lays the foundation of a modern expert system for the raw materials domain** with notably an ontology-based Dynamic Decision Graph and a database of methods and tools used in mineral intelligence, in geology and mining. In practice, the system helps a user to gain understanding and to find information on almost any topic related to mineral resources and the whole supply chain from prospecting to recycling, taking into account the environmental, political and social dimensions.

GeoSciML: The **GeoSciML data model** is an XML–based (conversion of a **UML** package) data transfer standard for the exchange of digital geoscientific information. It accommodates the representation and description of features typically found on geological maps, as well as being extensible to other geoscience data such as drilling, sampling, and analytical data (see: http://www.geosciml.org/).

Harvesting system: The **Minerals4EU (Central) Harvesting System** including the database periodically refreshes the information available about mineral resources by requesting data from the data providers using INSPIRE compliant Web services (**WFS**). This DB is structured in such a way that a large part exactly reflects the **INSPIRE** Mineral Resources (MR) data model, but it also includes the ProSUM mining waste modifications. During the harvesting phase the data that is received is checked whether codelists conform to the INSPIRE registry codelist values and other data have the correct format (e.g., dates, numbers…). The **Minerals4EU Harvesting DB** is hosted by the Geological Survey of Slovenia (GeoZS) and connected to the **Diffusion DB** using **SQL** scripts. This Harvesting DB delivers data related to primary mineral resources and mining wastes. **The ProSUM Harvesting DB** has been built using the **ProSUM unified data model**, and is dedicated to the urban mine (WEEE, ELV & BATT). It is hosted by the Geological Survey of Danemark and Greenland (GEUS) and data is currently extracted from Excel sheets that have a standardized format (**portrayals**) provided by the different ProSUM workpackages.

INSPIRE: The INSPIRE directive lays down a general framework for a Spatial Data Infrastructure (SDI) for the purposes of European Community environmental policies and policies or activities which may have an impact on the environment. The INSPIRE Directive entered into force on 15 May 2007. INSPIRE is based on the infrastructures for spatial information established and operated by the Member States of the European Union. The directive addresses 34 spatial data themes needed for environmental applications, among which Mineral Resources and Geology. To ensure that the spatial data infrastructures of the Member States are compatible and usable in a community and transboundary context, the INSPIRE Directive requires that additional legislation or common Implementing Rules (IR) are adopted for a number of specific areas (metadata, interoperability of spatial data sets and services, network services, data and service sharing and monitoring and reporting). These are published either as Commission Regulations or as Decisions. See: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007L0002&from=EN

INSPIRE MR data model: This is the European approved data model for mineral resources (MR), including both primary and secondary (i.e., Mining wastes) resources. However, mining wastes do not belong to the core part of this data model, being only an extension. One objective of the ProSUM project has been to improve and extend the mining wastes part of the INSPIRE MR data model.

(http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_MR_v3.0.pdf) (See deliverable D5.5). The mineral resources data model used in ProSUM is the M4EU data model (Minerals4EU project) directly derived (with **GeoSciML** extensions for geology) from the INSPIRE MR data model.

JSON (**JavaScript Object Notation**) is a data-interchange format. Although not a strict subset, JSON closely resembles a subset of JavaScript syntax. Though many programming languages support JSON, JSON is especially useful for JavaScript-based apps, including websites and browser extensions.

SQL (script): **SQL** or **Structured Query Language** is a special-purpose programming language designed for managing data held in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDSMS).

Structured data refers to any data that resides in a fixed field within a record or file. This includes data contained in relational databases and spreadsheets. Structured data first depends on creating a **data model**, i.e., a model of the types of business data that will be recorded and how they will be stored, processed and accessed. This includes defining what fields of data will be stored and how that data will be stored: data type (numeric, currency, alphabetic, name, date, address) and any restrictions on the data input (number of characters; restricted to certain terms…). Structured data has the advantage of being easily entered, stored, queried and analyzed.

ToolStack refers to a set of tools/softwares needed to perform a complex task such that no additional tools/softwares are needed to support this task.

UML, the **Unified Modeling Language** is a standardized general-purpose modeling language in the field of software engineering. It is a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system.

Unstructured Data (or **unstructured information**) refers to information that either does not have a pre-defined data model or is not organized in a pre-defined manner. Unstructured information is typically text-heavy, but may contain data such as dates, numbers, and facts as well. This results in irregularities and ambiguities that make it difficult to understand using traditional programs as compared to data stored in fielded form in databases or annotated (semantically tagged) in documents.

Web service: is defined by the World Wide Web Consortium (W3C - https://www.w3.org/) as 'a software system designed to support interoperable machine-to-machine interaction over a network'. Several types of web services are used by the EU-UMKP such as Web Feature Services (**WFS**) allowing the transfer of data, and Web Map Services (**WMS**) allowing the visualization of maps.

DELIVERY REPORT

1 – Introduction

Data on primary and secondary raw materials are available in Europe, but scattered amongst a variety of institutions including government agencies, universities, NGOs and industry. These data are often stored in databases (DBs) with their own design/architecture and vocabulary, making any attempt of merging them difficult and time consuming. The problems regarding availability, quality, organization, accessibility and sharing of these structured data (i.e., coming from DBs) are common to a large number of policies and are experienced across the various levels of public authorities in Europe. 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[3], but its implementation in the Member states is just starting and achieving those objectives remains a major challenge.

The development of the Urban Mining Knowledge Data Platform (EU-UMKDP) aims to give a simplified, user-friendly and efficient access to all available and new data related to secondary raw materials (Waste Electric and Electronic Equipment [WEEE], End-of-Life Vehicules [ELV], spent batteries [BATT] and mining wastes [MW]) from national geological surveys, scientific institutes and universities, relevant industries and professional organizations, as well as from former European projects such as ProMine[4,5] (providing information on both mineral deposits and anthropogenic concentrations resulting from mining and downstream activities) and EuroGeoSource[6,7] (providing information on energy and mineral resources, extraction locations, production, reserves). The system is also designed to accommodate and manage semi- and non-structured data, including reports and syntheses and statistics in the form of graph charts, time-series…, maps, images and videos in various formats.

The EU-UMKDP is thus designed to provide the end user with a seamless access to the whole value chain from mining, extraction and beneficiation of ore, characterization of products and CRM parameters, of stocks and flows of products, of waste streams and deposits arising from end-of-life products and mining wastes. It will combine all spatial and non-spatial pertinent information in a single reference system continuously updated thanks to its distributed architecture.

2 – Guiding principles for building the EU-UMKDP

The principles of the EU-UMKDP architecture follow the Implementing rules of the European INSPIRE Directive[3] to make data and services interoperable across Europe over a distributed infrastructure. The objective might be to extend the existing INSPIRE MR data model and to add a new thematic related to Urban Mine.

The role of the EU-UMKDP is to complete the EU-MKDP with the incorporation of information (structured, semi- and non-structured) for WEEE, ELV and BATT. Mining wastes data provided by geological surveys and former European projects such as ProMine include data delivery through INSPIRE compliant web services and according to INSPIRE data models. The data related to mining wastes, which have been augmented within ProSUM, are served by the EU-MKDP in order to keep mining and direct downstream activities linked **but are also accessible by the new EU-UMKDP** (see Figure 2) in order **to have the capability to combine information on all types of wastes**.

The open architecture defines various components connected together to build the EU-UMKDP, taking into account state-of-the-art developments to enhance the performance, stability, sustainability and user friendliness of the system. These components are connected using standardized interfaces.

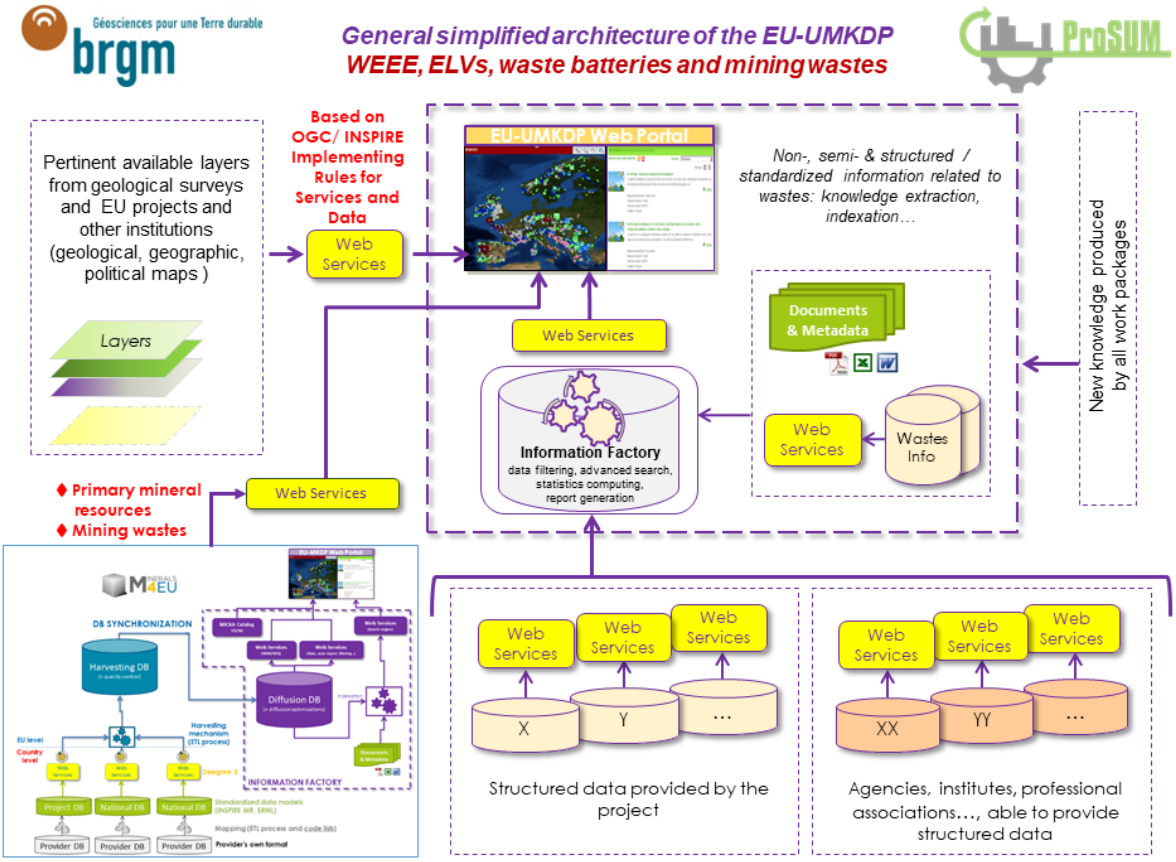


Figure 2 - Simplified architecture of the EU-UMKDP (EU-Urban Mining Knowledge Data Platform), and its link with the Minerals4EU Knowledge Data Platform (the EU-MKDP).

Figure 2Figure 2 explains how the system is designed to accommodate both structured data from existing databases (data produced by external providers and/or synthesised by project partners) and semi- or non-structured information produced by ProSUM work packages (reports, notes, publications, presentations…). The figure also shows the link with the EU-MKDP (Minerals4EU Knowledge Data Platform) which provides the EU-UMKDP portal with information related to mining wastes (ProSUM having contributed to improve the quality and completeness of mining wastes data).

2.1 - EU-UMKDP data & knowledge

The EU-UMKDP relies on the following data (or data type):

- Structured data on primary (mineral and ore deposits) and secondary resources (mining wastes) from databases provided by geological surveys through EU-FP7 Minerals4EU and ongoing EURare projects. These two projects reuse the advances made in EU-FP7 ProMine[4,5] and EuroGeoSource[6,7] projects, respectively in terms of data collection/collation and architecture. The Minerals4EU and EURare platforms share the same harvesting system (see Figure 1), but the 'operational link' with the ProSUM KDP is established through the Minerals4EU EU-MKDP (see Figure 2). Data is delivered through harmonized datasets according to the INSPIRE (plus extensions) data model[8] and ERML v.2 data model[9,10,11,12,13]. Both the EU-MKDP and the EU-UMKDP, access to the Central Diffusion Database through web services (see below).
- Structured data on secondary raw materials from the urban mine:
  (i) generated or collected (e.g., national and pan-European databases stored at Eurostat [including Member State data returns for the ELV Directive, Battery Directive, WEEE Directive, Waste Statistics Regulation, International Trade Statistics, Prodcom Statistics]) and collated by ProSUM work packages and then harmonized and homogenized and stored in dedicated Excel portrayals which are then uploaded in the ProSUM Harvesting DB.
  (ii) provided by external agencies/institutes/organizations like EUROSTAT. This data (related to EEE products and stocks) is first extracted from the provider' DB, transformed using correlation tables and a specific data model (the UNU-WEEE data model) to be stored in the UNU database (developed for UNU's own purpose research activities). The UNU-WEEE data model, although being not so 'distant', is nevertheless different from the ProSUM unified model. The data is therefore transformed into a dedicated portrayal before being ulpoaded in the ProSUM Harvesting DB which implements the ProSUM unified datamodel (see Figure 3).
- Information in various formats (i.e., semi- and non-structured) and delivered during the project by ProSUM work packages (e.g., statistics analyses, analyses of supply and demand in the EU, stocks, flows, including trends in what products are put on the market and their composition, analyses of the composition of waste products and wastes arising from pre-processing…) is stored in the ProSUM Knowledge Base.
- All this data, whatever its origin, is public[1]. At this stage, no dedicated system for managing confidential data/information is scheduled.

[1] Sensible data provided to the project by industrial partners has been used for certain calculations. It is thus 'embedded', and is not accessible in its original format.

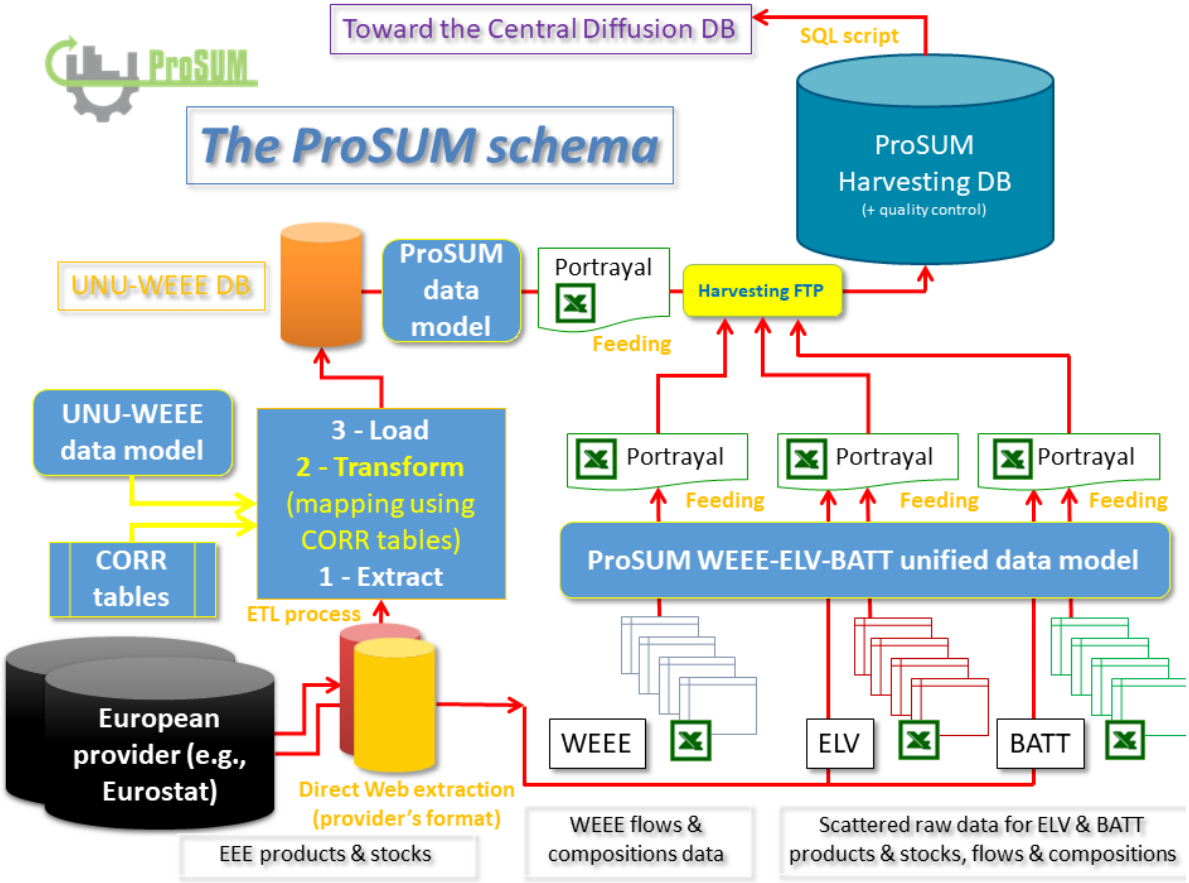


Figure 3 – Representation of the data flow in the ProSUM Project schema. Note that most of the scattered data which are collected and collated, are homogenized using Excel spreadsheets, so-called 'Project Excel databases'.

2.2 - EU-UMKDP components

The main components of the EU-UMKDP architecture are the followings:

- Web Feature Services (WFS) for the delivery of the structured data on primary and secondary mineral resources (mining wastes). These services are INSPIRE compliant. They are delivered to the EU-UMKDP using the same web services as those used for the EU-MKDP.
- Services (possibly FTP – see the note at the end of section 3.3) associated to upload ProSUM-based Excel portrayals for the delivery of the structured data on secondary raw materials related to the Urban mine (see Figure 3).
- A dedicated Harvesting Database connected through SQL scripts to the Central Diffusion Database shared by all platforms (see also Figure 1).
- A metadata catalogue (CS/W) for the delivery of relevant external spatial layers, e.g., geological, geographical, political and land use, provided as interoperable services that will be registered within the project. This metadata catalogue is accessible via an INSPIRE discovery service.

- A map viewer for the geographic representation of all spatial data. At this stage, only mining wastes will be plotted within this map viewer. For all other information related to WEEE, ELV and BATT, this is statistical data without geographical reference (except the country).
- A semi- or non-structured information management system. This system will provide the user with the possibility to add semi- or non-structured document to the KDP and its Knowledge Base by creating some metadata records (using Dublin Core type, see below). These documents and their metadata records will then 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 for each facet. The way to display the list of selected documents will take into account these three facets.
- A search engine 'tackling' all data, layers and documents delivered by the project and all external pertinent documents.

3 – The EU-UMKPD (EU-Urban Mining Knowledge Data Platform)

In order to increase the value of the work done in former EU-FP7 projects, and also to minimize new developments, the EU-UMKDP reuses the developments made in the framework of EU-FP7 projects EURare (IKMS) and Minerals4EU (EU-MKDP) for both the harvesting and the diffusion systems.

3.1 – Detailed architecture and functioning principles

The EU-UMKDP architecture is made up of web services,, databases that are implementations of data models, Extract-Transform-and-Load (ETL) processes, and catalogues. A code list classifications has been developed to populate the databases for the urban mine. This system takes into account how the data model will function within the architecture of the EU-UMKDP and the granularity of the data we want to present.

A data model (see also the glossary) is the representation of a domain, of its objects (called 'features') with their intrinsic properties, of the links/interactions between these objects, with defined vocabularies (code lists) as parts for describing these properties. Data models are an agreed representation of a domain, which will 're-organise' the data provided to make them harvestable by the system. This is called the ETL process and the preparation of the Web Feature Services (WFS). This enables the 'digestion' of different data formats and the creation and linking of different databases, used internally as well as externally.

The harvesting system retrieves the harmonised data. The role of the Harvesting database is to store this harmonized data, to enable quality control and to warn when abnormal or suspect figures are detected. The Harvesting and Diffusion databases are synchronised using SQL scripts. The Diffusion database for its part optimises data from the data models to improve the responsiveness (speed) of the portal accessing the EU-UMKDP, for example by storing the results of pre-defined queries into separate tables.
Error! Reference source not found. shows that information can either be extracted from a Provider DB (e.g., EUROSTAT) or compiled in Excel spreadsheets. This information can be then transformed via correlation tables and possibly a dedicated data model (here the UNU-WEEE data model

developed for UNU's own purpose research activities[2]), before being uploaded to the Central Harvesting database. This sequence of operations is a type of ETL process, Extract (data from the Provider DB), Transform (data to fill [or fit with] the data models) and Load (data into the harmonised database/portrayal). Once this data is stored in their harmonised database/portrayal, the Harvesting System is able to retrieve it and store it in the Harvesting Database.

This ensures that data with different sources related to product and stocks, and flows and composition of the secondary raw materials are "described" in the same way, with the same fields, using the same vocabularies.

It should be noted here that, technically, the correlation tables will need to be updated yearly as new products are regularly coming on the market and/or in external databases used for data preparation.For instance the Combined Nomenclature (CN) can have additions or expirations. The ETL process will be as much as possible automated in order to minimize human intervention, also having in mind the maintenance of the system in the future, after the end of the project.

One of the major advances made by ProSUM is that **by using a harmonised classification system on which the data model is based, heterogeneous data are transformed into a harmonized format, ultimately allowing the linking multiple data sources and re-processing of new, updated and homogeneous information**. In the context of the projet, this means that data collected in 'Provider/Project Excel Databases' is consolidated in the Harvesting Database and is then forwarded to the Diffusion Database to feed the various applications running on top of the EU-UMKDP.

3.2 – Data models and vocabularies

ProSUM 'reuses' for the mining wastes (MW) part the ToolStack from Minerals4EU and EURare because it fulfils the need for involved partners. The UML and DB modelling uses Enterprise Architect licensed version (Sparx system), but a free viewer is available. The other ToolStack elements are open source software but specific versions are used for compliance purposes:

- Apache Tomcat version 8.0.15
- Deegree version 3.3.13
- PostgreSQL version 9.2.9-1
- PostGIS 2.0.1-1
- GeoKettle version 2.6-r192

All open source softwares can be found in the Minerals4EU repository at:
http://data.geus.dk/svn/m4eu/

The modelling phases are the following:

1. Model /update the DB UML diagram;
2. Model/update the DB model diagrams;
3. Create the DDL data for the DB model;
4. Update the WFS schemas for the Deegree;
5. Create/update code list scripts;
6. Send out the update scripts to partners.

[2] Note that with the 'UNU step', ProSUM benefits from the harmonisation work done by UNU with correlation tables, but has nevertheless to re-map the data over the ProSUM unified data model as the UNU-WEEE data model is different.

For the modelling and the extension of the mining wastes part, ProSUM has benefitted from the experience acquired in the Minerals4EU and EURare INSPIRE UML model and DB model development. This modelling part has also implied scheduling updates appropriate to all Minerals4EU and EURare partners.

In paralel, ProSUM has built up a new UML and DB model for secondary raw materials from the urban mine. This database is a stand-alone (harvesting) DB which thus doesn't imply any update for Minerals4EU and EURare partners.

3.2.1 – Data model and vocabularies for the urban mine

Figure 4

Figure 4 summarizes the main parameters the model must take into account from product characterization to stocks and flows management and to waste characterization, knowing that the parameters are given by country and that they are all continuously evolving with time…

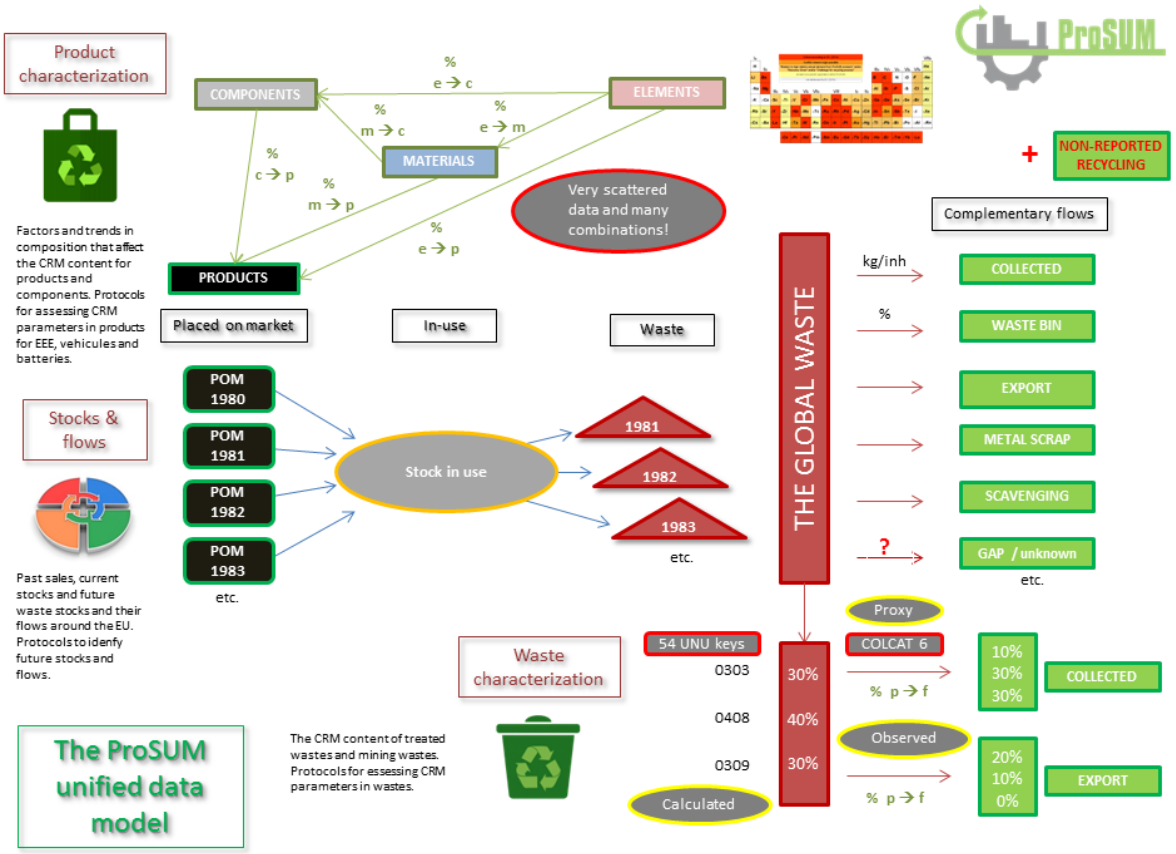


Figure 4 - Some of the main parameters managed by the ProSUM unified data model

A new model has been developed in this project that unifies the data aspects of the three urban product waste types, i.e., WEEE, ELV and BATT (see Figure 5, also available at a larger format here).

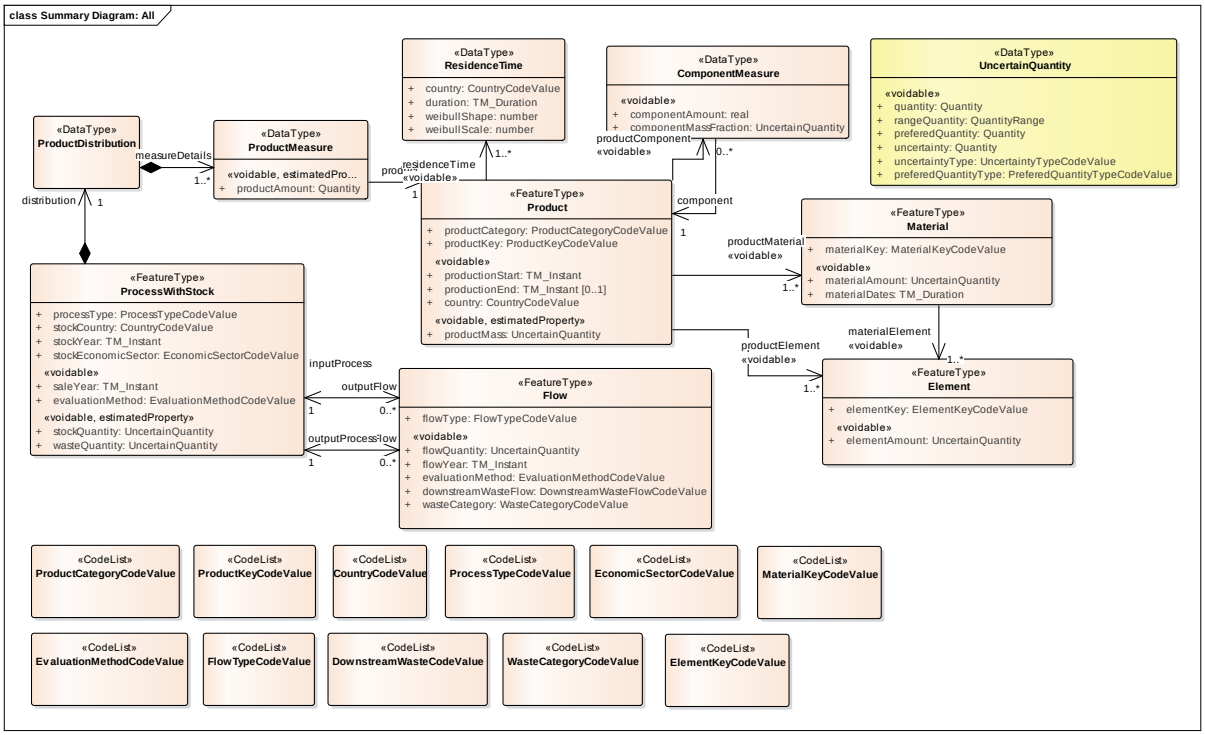


Figure 5 UML class diagram describing the unified conceptual data model for WEEE, ELV and BATT.

This new unified data model allows describing products of the three wastes groups in detail with their special features (like the residence time which is needed to know the proportion of this product that will become waste for a certain date), and also with their composition in smaller components, materials or elements. These components can then be detailed with their own composition, in the same way the materials used can also be described with the component that are used to made them.

The entire life of the products can be described with the different flows and stocks the products will follow. The two concepts of ProcessWithStock and Flows are interlinked and can be used in combination or in a separate way as it is requested by the different actors of the domain who can represent the life of the products by moving stocks or flows between these stocks.

A PostgreSQL database implementation of this model has been developed and datasets provided by work packages (WP) 2, 3 and 4 have been uploaded and inserted in the database tables. The vocabulary that has been created in WP2 to WP4 is stored in 13 code lists with 9 currently containining values (

Table 1).

Code list name	Source
CountryCodeValue	ISO-3 values available

ElementKeyCodeValue	
EconomicSectorCodeValue	**empty, developed for future use**
DownstreamWasteCodeValue	
FlowTypeCodeValue	**empty, developed for future use**
MaterialKeyCodeValue	
PreferredQuantityTypeCodeValue	
ProductCategoryCodeValue	
ProductKeyCodeValue	
ProcessTypeCodeValue	
UncertaintyTypeCodeValue	
UnitOfMeasureCodeValue	
WasteCategoryCodeValue	**empty, developed for future use**

Table 1: Code list and sources for WEEE, ELV and BATT.

These code lists values harmonize the data for WEEE, ELV and BATT that are available in Europe. A functionality has been created so that these data can be uploaded to the database system. This functionality can be extended or modified according to future needs for retrieving urban mine data.

Note: The ProSUM unified data model and associated code lists are described in detail in Heijboer et al. (2017)[2].

3.2.2 – INSPIRE /ERML model and vocabularies for mining wastes

The Minerals4EU data model, which extends the INSPIRE model, is based on EarthResourceML version 2.0 (ERML) and allows associating mining waste information with specific minerals deposits. However, it only allows storing information about the amount, grade and density of one waste material. It does not include data on the presence, types and amounts of critical raw materials (CRM commodities) which is necessary in the ProSUM project.

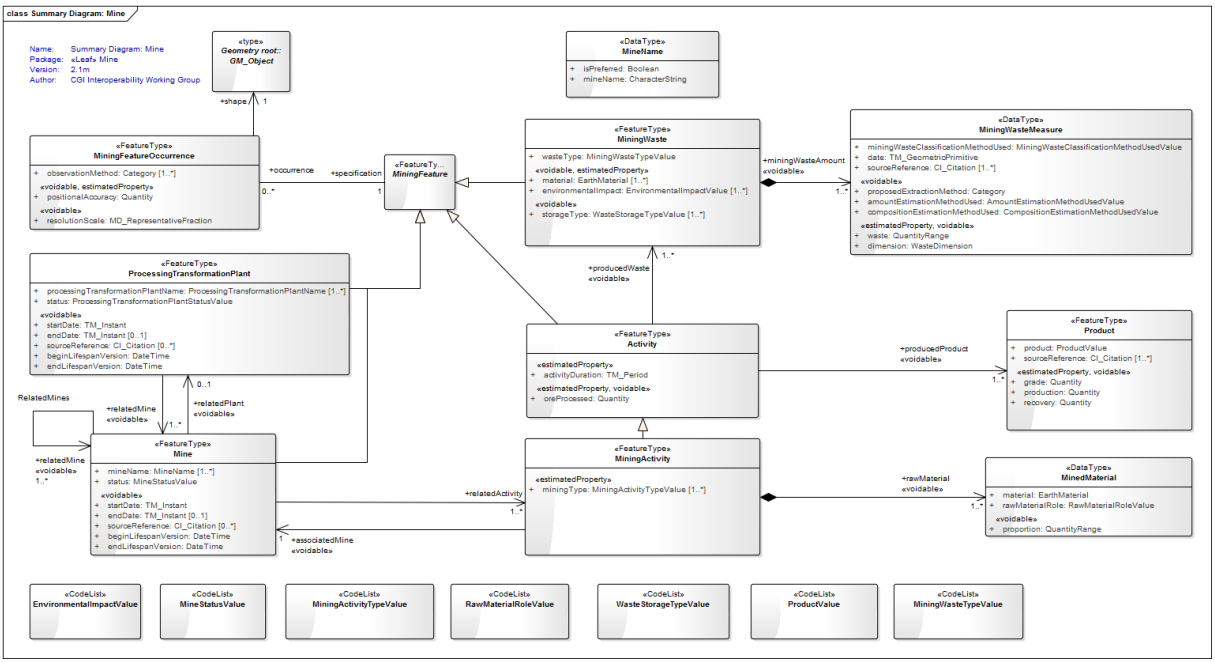


Figure 6 - Summary Diagram: Mine

The MiningWaste feature is therefore linked to a new MiningWasteMeasure concept which is quite similar to OreMeasure (used for primary resources) but adapted to the specifics of mining waste. For example a **WasteDimension** type has been created to handle density, volume and area of the MiningWasteMeasure and the **miningWasteClassificationMethodUsed**, **proposedExtractionMethod**, **amountEstimationMethodUsed**, **compositionEstimationMethodUsed** attributes have been added. This new WasteMeasure is linked to the CommodityMeasure concept used for Earth Resources, similar to OreMeasure. In addition to this, a correction has been made to the link between CommodityMeasure and Commodity (cardinality was 1 to many, but is reduced to 1).

In the ERML V2.0 data model, a MiningWaste feature was linked to a MiningActivity feature. In the modified version MiningWaste is linked to the Activity Feature (Figure 6). Both MiningActivity and ProcessingTransformationActivity (see below) are derived from Activity. MiningActivity itself can be associated to 0 or 1 MineralOccurrence, whereas there can be several MiningActivities associated with a MineralOccurence and each MiningActivity can generate one or many MiningWastes. A Mining Waste is linked to a MiningActivity but can be independent of a Mineral Occurrence if no information about a MineralOccurrence is known. The location of a MiningWaste is represented by the geometry property of the MiningFeatureOccurrence that is linked to the MiningWaste. It is not dependent on other features.

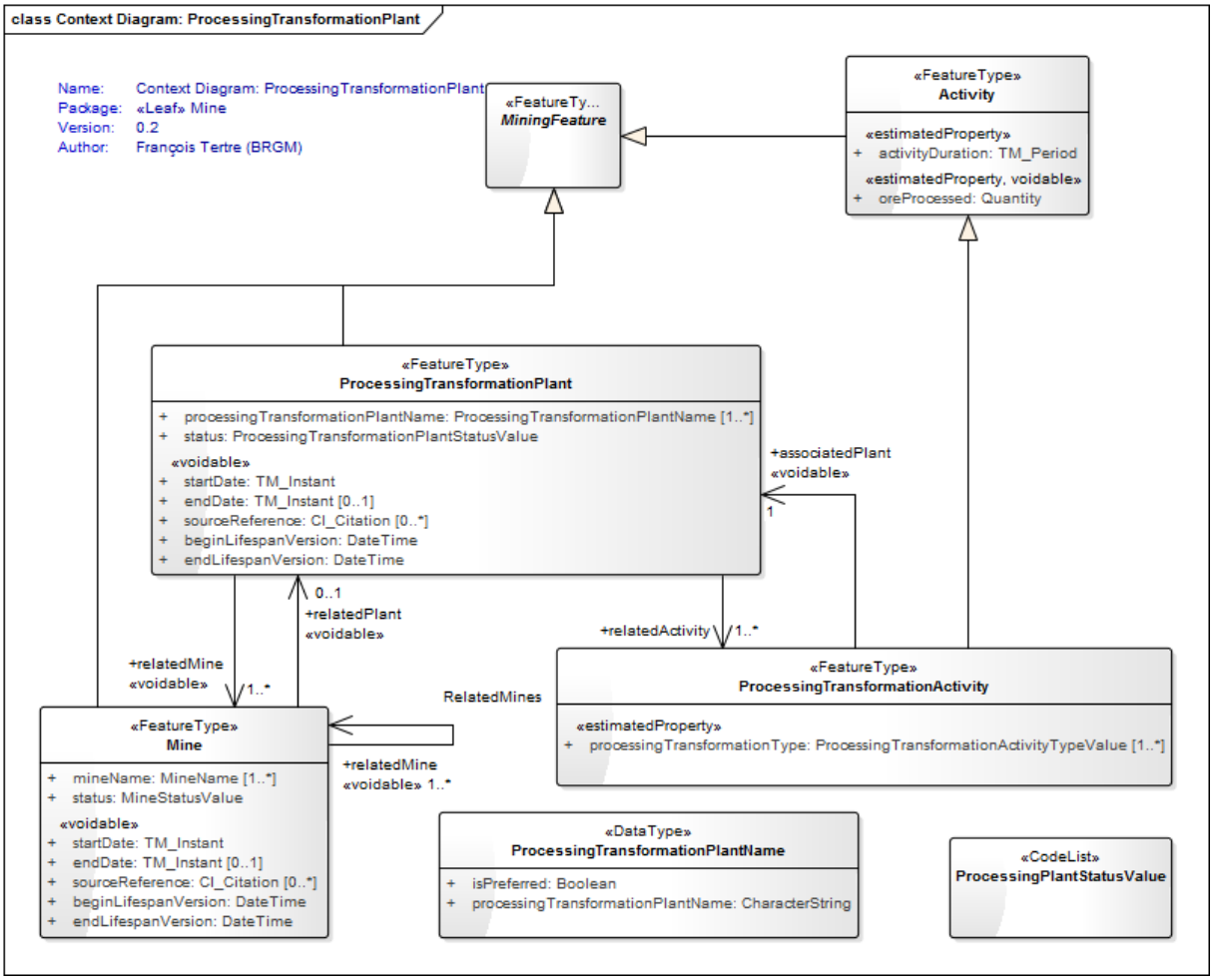


Figure 7 - Summary Diagram: ProcessingTransformationPlant

In addition to these changes a new concept **ProcessingTransformationPlant** has been introduced in the model introduced (Figure 7). It represents the plant that processes the mined material coming from one or more mines, and/or the smelter that transforms mine products which can be as simple as a (very) rich ore or a concentrate. This concept allows a more detailed description of how mined material was processed and how this relates to wastes (tailings). This description includes downstream activities like smelting and captures specific wastes such as slags and fly ashes. It also makes it easier to extend the data model to primary transformation industries in general, notably those related to industrial rocks. A ProcessingTransformationPlant has one or many **ProcessingTransformationActivities**. The **ProcessingTransformationActivity** is an extension of the Activity Feature (as described before) and represents a period in which mined material was processed in a specific way. **The ProcessingTransformationActivity** consumes either MinedMaterial that is extracted from the Mine or Products of a mining activity (e.g., a concentrate). This new concept is explained in Figure 8.

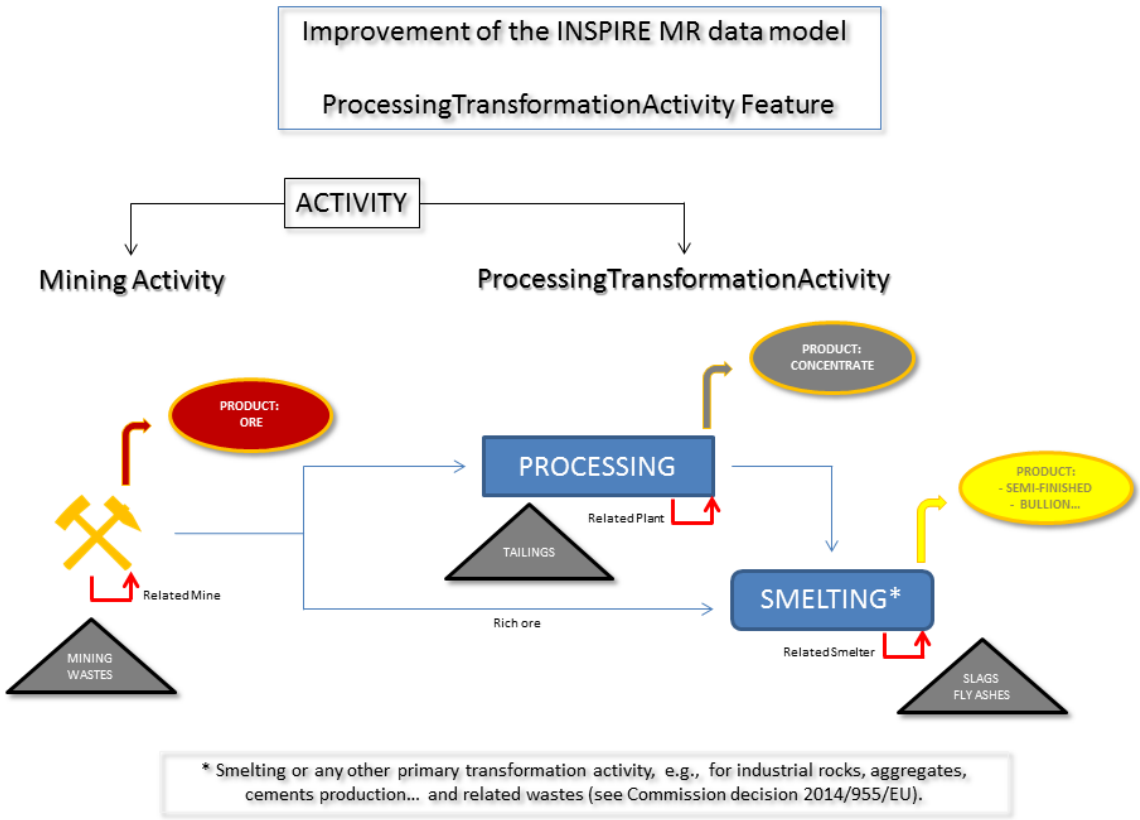


Figure 8 - The ProcessingTransformationActivity Feature

Thus, in summary:

For mining waste the Minerals4EU model was modified and extended in such a way that it still delivers INSPIRE (v3.0) compliant mineral resource Web feature services (Figure 9 – see a larger version here), defined in the context of the ProSUM project.

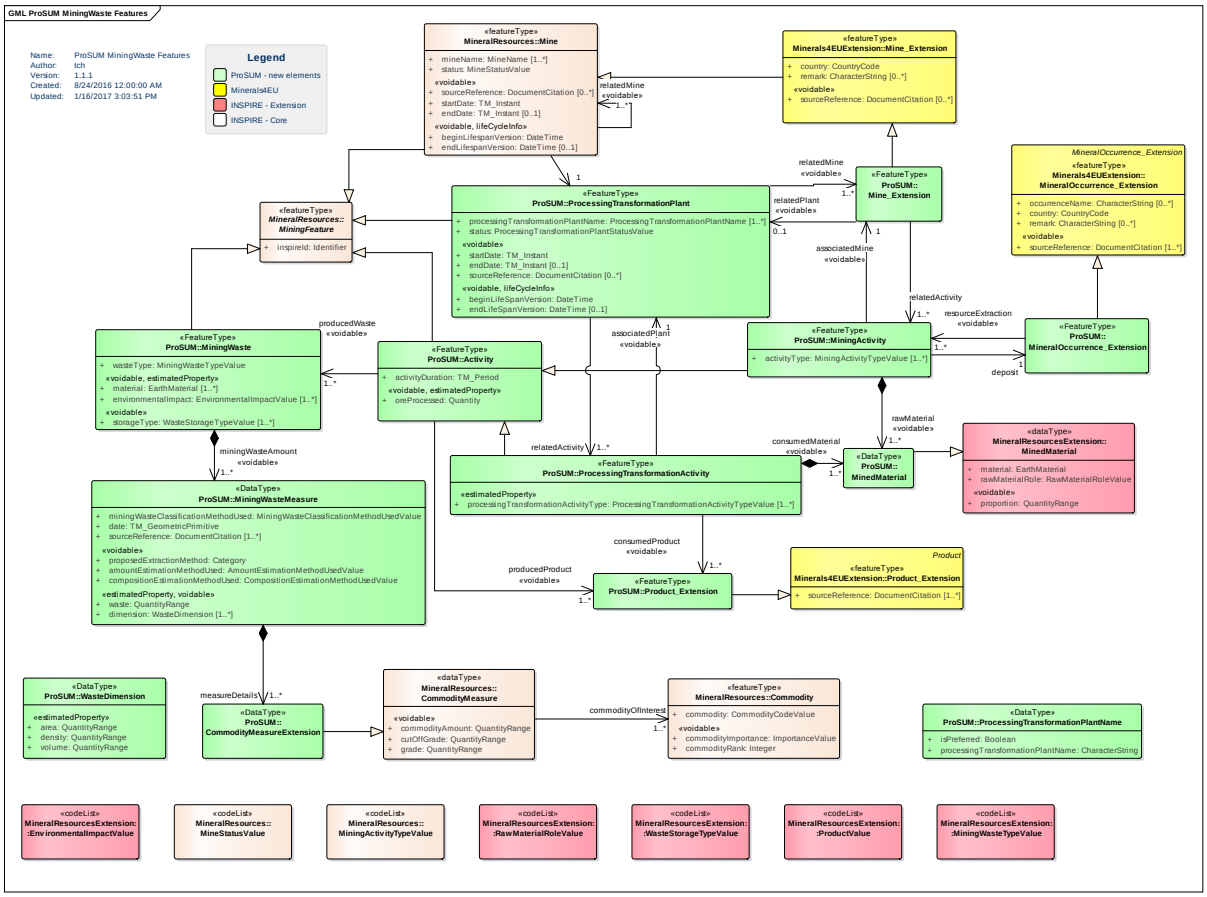


Figure 9 - UML class diagram summarizing ProSUM modifications to the M4EU conceptual data model (green)

A new version of the **MiningWaste** and **MiningActivity** classes have been developed and 2 completely new classes have also been introduced: **ProcessingTransformationActivity** and **ProcessingTransformationPlant**. These additions make it possible to describe mining activity and mining waste in greater detail and allow information about the processing of mining products or waste to be provided similar to other aspects of mineral resource information. It is therefore an accurate description of the material flow related to mining. The **MiningWasteMeasure** class (which is linked to the **Commodity** class and the **WasteDimension** class) allows for the description of amounts of elements of interest present in a mining waste dumps or tailings.

The changes described may in turn be incorporated into a future version of the INSPIRE Mineral Resource (MR) data specification and application schema. It is envisaged that Geological Surveys who are beneficiaries in the ProSUM project and which have collected mining waste data will deliver data in the form described. Geological Surveys that are not a part of the ProSUM project can choose to use the database scripts and Web Feature Services setup that have been developed in this project by downloading them from the repository at http://data.geus.dk/svn/m4eu. They will be encouraged to do this in the frame of a new proposal called 'ORAMA', which has been accepted in July 2017 and which aims at improving/optimizing European datasets related to both primary and secondary resources.

The changes of the Mining Wastes part of the ERML data model are also accompanied by the development of new code lists. Two new code lists related to the methods for estimating the

amount (**AmountEstimationMethodType**) and the composition of mining waste (**CompositionEstimationMethodType**) have been proposed and already submitted to the IUGS/CGI/GTWG[3] for international (worldwide) acceptance and then to the INSPIRE validation process.

Note: The new code lists related to mining wastes are described in Heijboer et al. (2017)[2]**.** The other ones can be consulted in Schubert et al. (2014)[15].

3.3 - The data provider system

One particularity of the ProSUM Urban Mining Knowledge Data Platform compared to the IKMS or the EU-MKDP is that there is no National data provider. The only National data providers involved are the European Geological Surveys which are serving Mining wastes related data, but this information is arriving in the EU-UKMKP through the EU-MKDP, thus not directly. The main ProSUM data provider is thus the project itself. These data have or may have an external source, e.g., Eurostat, but they are most of the time mixed/merged with other data coming from publications, reports, statistical studies… and then homogenised, structured and entered in portrayals (also called 'Project Excel Databases') after a mapping on the unified ProSUM data model. These portrayals are then sent to a FTP, and the files are harvested for the feeding of the ProSUM Harvesting DB. The WEEE schema is slightly different as the data extracted from the Eurostat DB are stored – as they are - in an intermediate DB which is then used to feed the UNU database built following the UNU-WEEE data model. This is this base which is then harvested by ProSUM for the products & stocks part, using this time a portrayal based on the ProSUM unified data model– see Figure 3 & Figure 10.

[3] http://www.cgi-iugs.org/tech_collaboration/geoscience_terminology_working_group.html. Proposal submitted during the IUGS/CGI Ispra meeting on October 29-30, 2015.

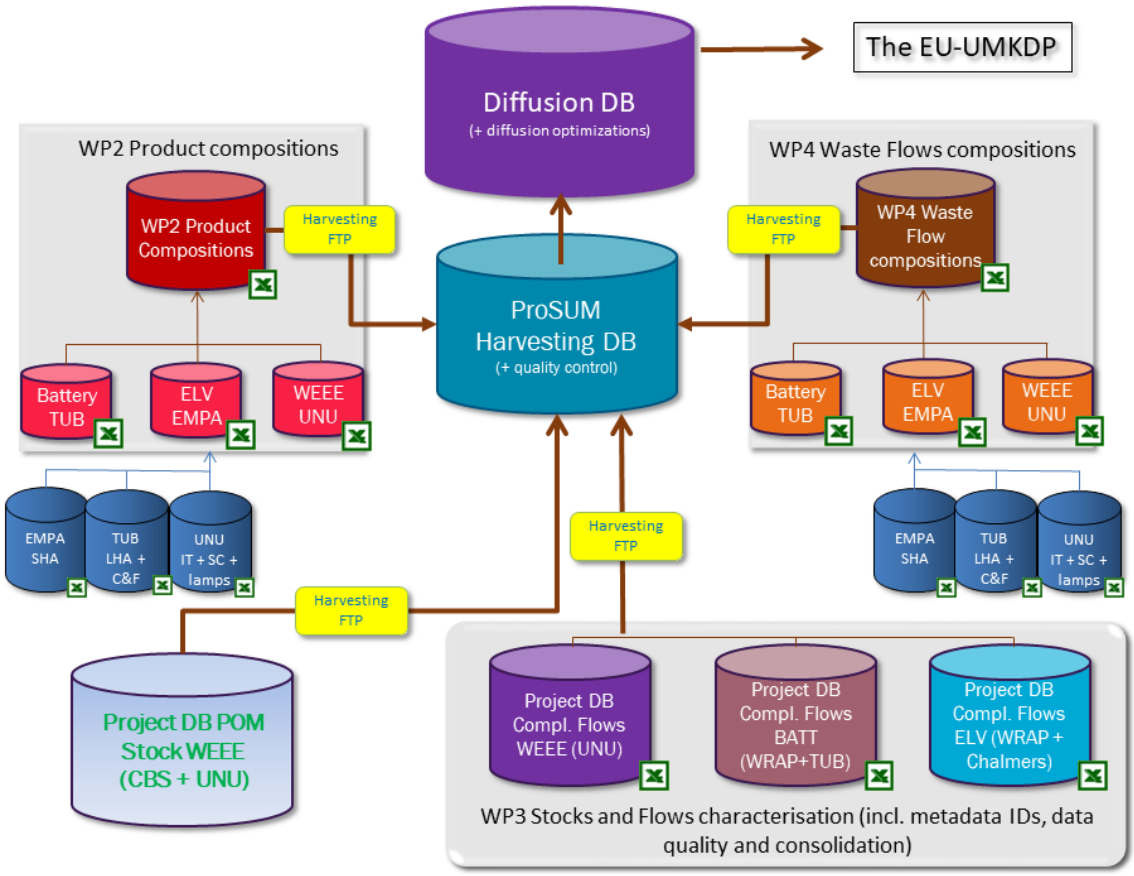


Figure 10 - Overview of the different databases, including the so-called 'Excel Project DBs', the different providers, and the way they will be harvested.

The future maintenance of the system will thus strongly relies on ProSUM partners for the Urban Mining part, hence the absolute necessity to set up a strategy, including several possible scenarii, with associated business plans and the research of synergies with other EU recently finalized or ongoing projects, like for example the Minerals4EU project, which will implement a Permanent Body (registered Foundation) for the maintenance of the EU-MKDP and the management of its network.

Note: The first version of the EU-UMKDP has not implemented a File Transfer Protocol (FTP) for uploading harmonized data in the ProSUM Harvesting DB. Data produced by the different work packages has been stored on a BRGM server using Seafile, an open source file sync and share software, for facilitating data exchange and corrections/modifications.

3.4 – The harvesting system

A harvesting system consists of a harvesting application and one or several databases.

The data about the Urban Mine are served as Excel Portrayal (Figure 11). Data from Excel Portrayal is then sent to the ProSUM Harvesting Database. **To set up a dedicated harvesting procedure like the one used in Minerals4EU for the EU-MKDP, is not justified here as the 'WEEE-ELV-BATT' data will be updated on a once-a-year basis**.

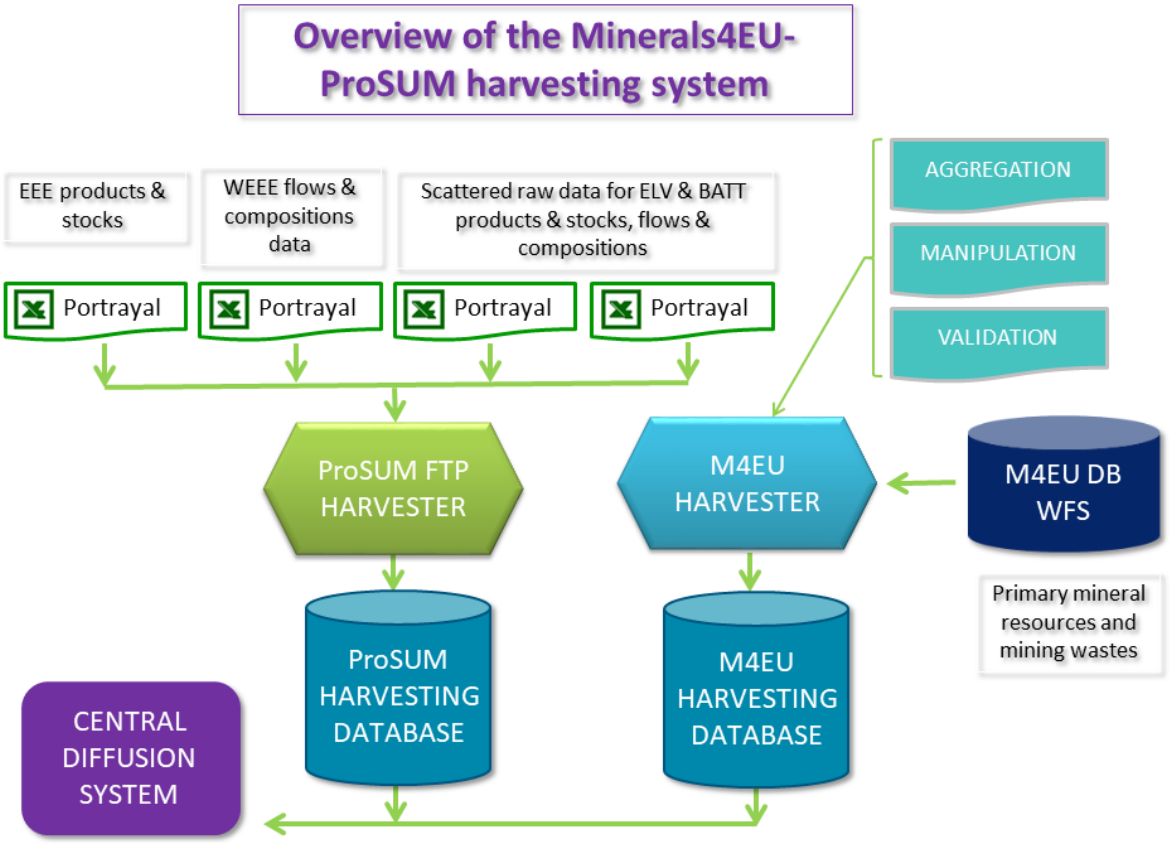


Figure 11 - Overview of the harvesting system

For the mining wastes, the Minerals4EU harvesting application retrieves WFS formatted data from each provider, harvests data from it, transforms it into proper form and then stores it in the Central Harvesting database. It was decided to develop the harvesting system using JAXB and with NetBeans.

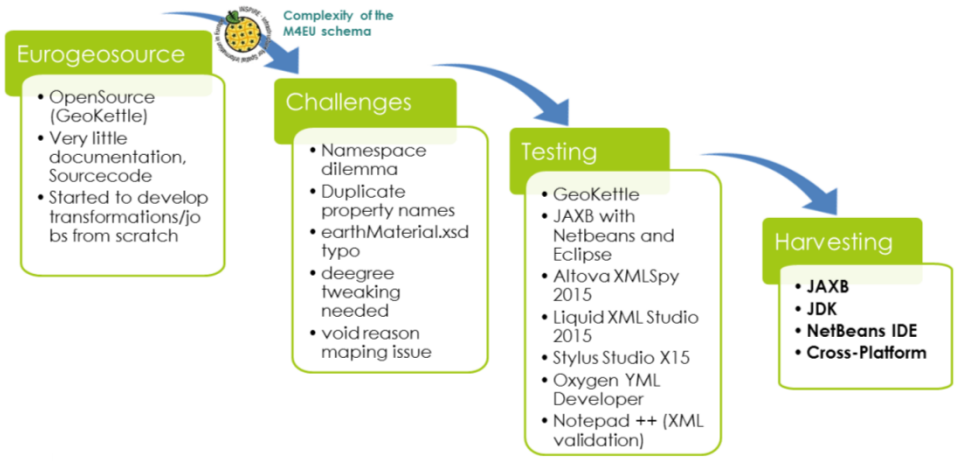


Figure 12 - The development of the Minerals4EU harvesting system and the challenges encountered

During the development of the Minerals4EU harvesting scripts it turned out that the GeoKettle ETL tool could not properly parse WFS 2.0 formatted XML files (see Figure 12). The biggest obstacle was parsing XPath from XML files. For this reason it could not properly navigate through the XML elements and attributes of the XML document.

JAXB (Java Architecture for XML Binding API) was tested both with NetBeans and Eclipse SDK using their own JAXB components and several different JAXB libraries. The JAXB offers easy access to XML documents from within applications written with Java programming language. In order to access and harvest data from XML document with JAXB, the first task is to generate several Java classes that represent the project schema. Then it is necessary to unmarshal the XML document into Java content objects. The Java content objects represent the content and organization of the XML document. It is then possible to store this java objects in the database. One of the advantages is that the objects contain all the relations to other objects. Aggregation and other manipulation of data in the harvesting system, if needed, will be performed by procedures within the harvesting application and PostgreSQL/PostGIS database. The harvesting database will have the same structure as the database (or portrayal) that each provider will setup. The delivery of data from the harvesting system to the diffusion system will be carried out by sending SQL update scripts.

Figure 13Figure 13 summarizes the detailed architecture of the whole ProSUM system.

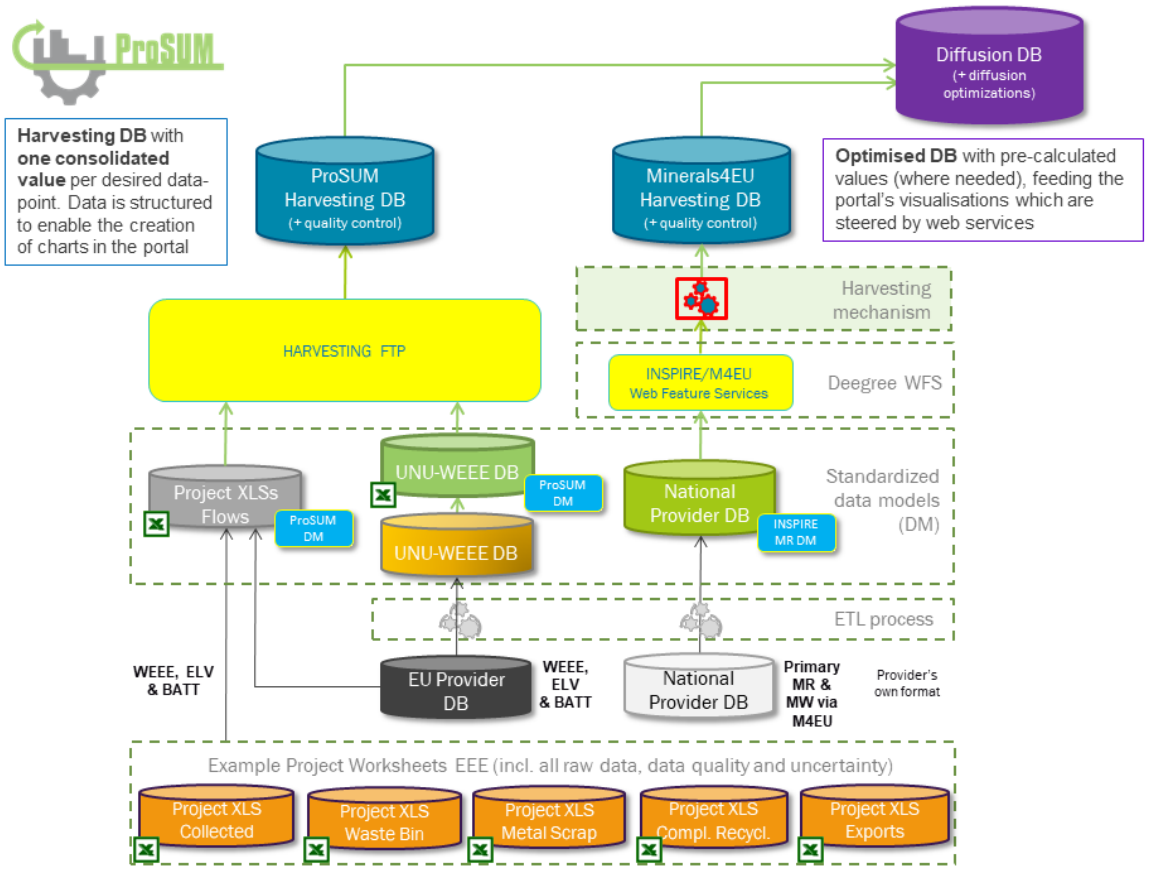


Figure 13 - The EU-UMKDP detailed harvesting system architecture

3.5 – The diffusion system

The diffusion system consists of a central database. This is a copy of the harvesting databases synchronized at regular intervals. The diffusion system also has components specialized for data delivery and services that carry out computations based on the stored data. These services include:

- Simple map visualizations using WMS (Web Map Services). Some parts of the content of the Diffusion Database are available in the Map Viewer of the Diffusion portal. These maps allow visualising the locations of the Mines, of the Processing Plants (ongoing), and of the Mining Wastes.
- Dedicated services: these services are described in a dedicated paragraph below.
- Data download: in some case (and depending on the access rights to the data) the data is or will be available for download so that it can be processed in other applications (e.g., desktop GIS…).
- Search facilities: these facilities allow the user searching in the whole database and documents.

The diffusion system proposes a CMS (Content Managing System) that offers the latest news of the system and a note to describe its general functioning (this note). This main access to the system is linked to the different parts, i.e., the Data Search, the Map Viewer and the Services (Figure 14).

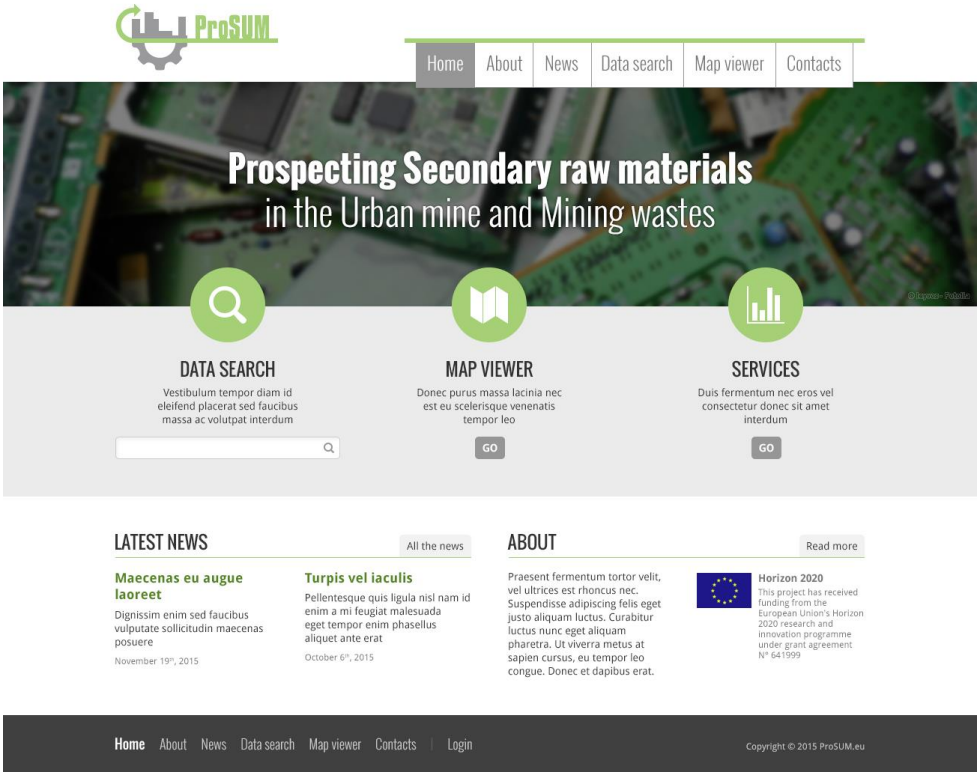


Figure 14 - Main access to the diffusion system (http://prosum.brgm-rec.fr/)

The Data Search (Figure 15Figure 15) allow searching in the Diffusion Database (containing data on mining wastes, ELV, BATT and on WEEE), but also in the public documents. This search can be simple (What, Where) or more complex if the user selects a data source (for a document, search in the creators of the document, in the title…; for Mining Waste, search according to the name of the related Mine, to the commodity…).

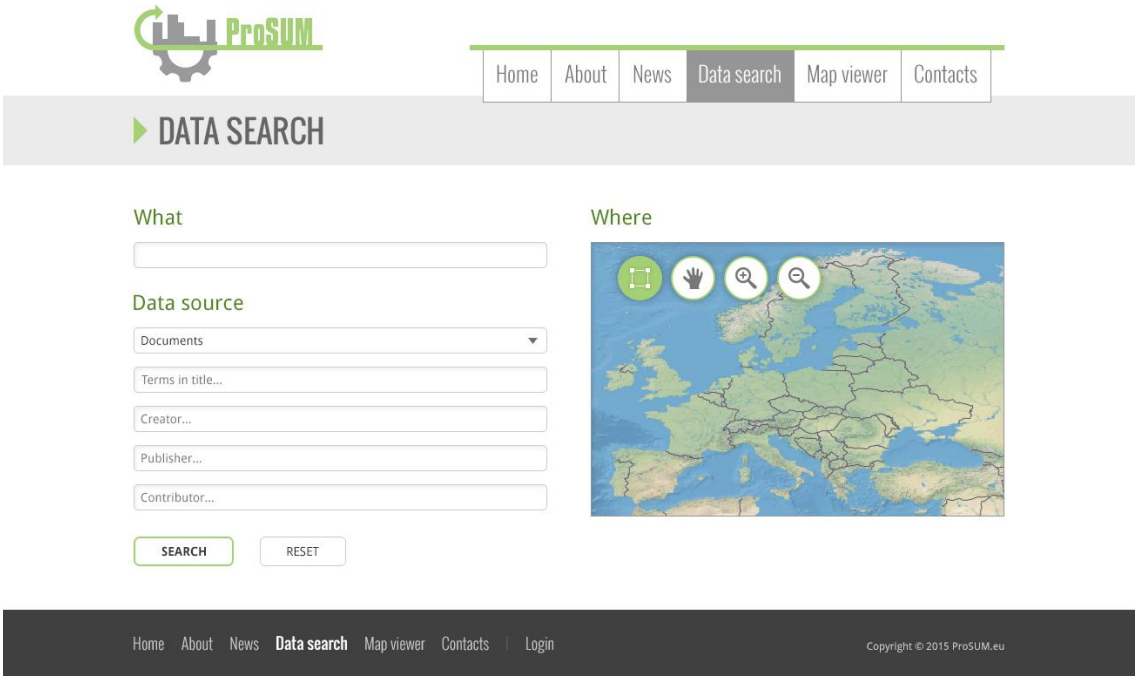


Figure 15 - Data search.

Then, the Map Viewer (Figure 16Figure 16 - Map Viewer) allows visualizing some parts of the Diffusion Database plotted on a map, and related to e.g., Mining Wastes, Mines or Processing Plants (ongoing work).

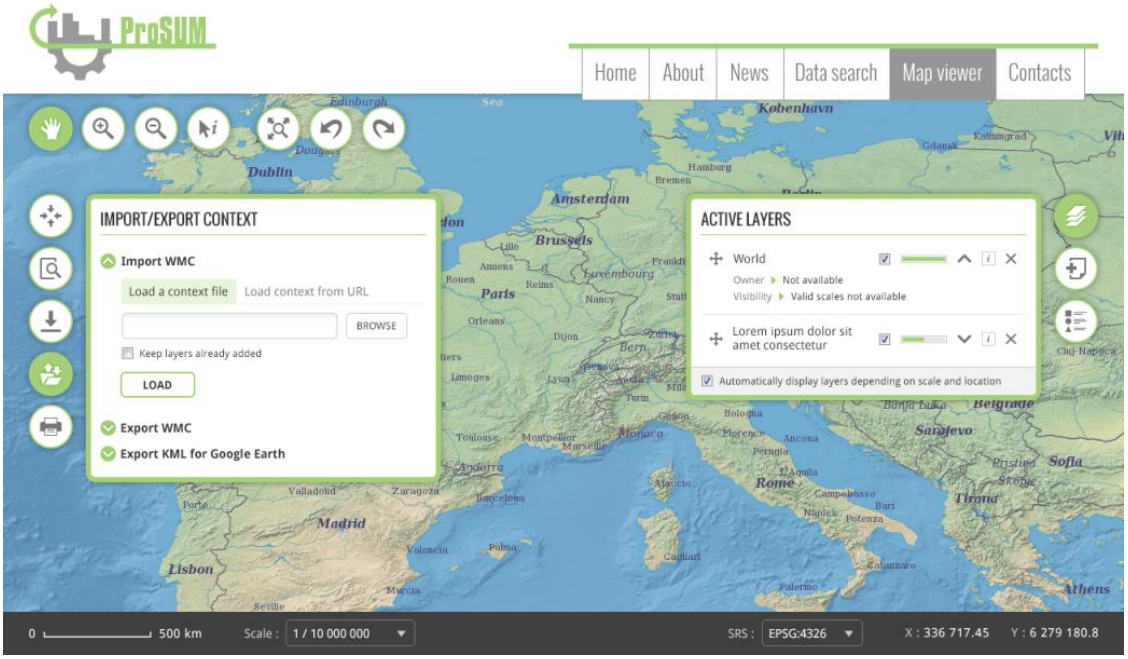


Figure 16 - Map Viewer

3.6 – The non-structured data and their metadata

Beside the Diffusion Database, an additional interface allows experts to feed the Knowledge Base with non-structured data. These non-structured data can be of different types (reports, synthesis notes, thesis…). This interface allows the expert to add a document (the document will be part of the EU-UMKDP and retrievable within the EU-UMKDP) or a link to an existing document in some place accessible via the Internet (the document will stay on its original server; the end user will have to follow the link to retrieve it). To be able to integrate it into the Knowledge Base, the expert needs to associate metadata to this document. These metadata are based on Dublin Core (http://dublincore.org/) with a specific extension (at least quality and accuracy of the data). Dublin Core Metadata Element Set has 15 elements covering the most important properties to describe a document (title, creator, subject, coverage – temporal or geographical… - see Table 2). These metadata are then be used in the search facilities to retrieve the documents.

Standard Dublin Core fields	
1Title	Input/Copy-paste full title of data-source *Specification of this descriptor is quite straightforward.*
2 Subject	Input/Copy-paste key-words of data-source including at least 1 from term-list 'Subject' (Products and Component composition/Stocks&Flows/Waste/General) *Copying the key-words of the data-source will allow a keyword search as well as compiling a list of all key-words in the bibliography for this field. Additionally, 1 term is deemed mandatory from the term-list 'subject' in order to enable WP-leaders to quick-scan bibliography entries of relevance to their WP.*
3 Description (Abstract)	Input/Copy-paste abstract or summary of data-source *Specification of this descriptor is quite straightforward. Copying the abstract or summary enables an expanded key-word search.*
4 Type	Select one or more from term-list 'Type' (text/image/etc.) *Specification of this descriptor is quite straightforward.*
5 Source	Specify origin of data-source *Specifying the source of the actual data on CRM's, is intended with this descriptor.*
6 Relation	Select from term-list 'Relation'(Original/Secondary/Compiled) *Specifying 'relation' is chosen to be defined as the internal relation to other documents of a specific data-source: Is it original data, or is based on other data-sources, directly or indirectly. Important references from the reference list may be inserted here as well.*
7 Coverage (Geo)	Select one or more country codes from term list ISO 3166 - alpha 3 for the geographical location of the actual data on CRM's *For ProSUM, geographic information for the data-source itself is irrelevant. It is deemed more useful if this descriptor is used for geographical information of the CRM's.*
8 Creator (Author)	Input/Copy-paste corresponding author *Specification of this descriptor is quite straightforward.*

9 Publisher/Journal /Institution	Input/Copy-paste publisher of data-source (leave blank if confidential) *Specification of this descriptor is quite straightforward.*
10 Contributor (Co-Authors)	Input/Copy-paste contributing authors *Specification of this descriptor is quite straightforward.*
11 Rights	Select from term-list (Copyright, conf., internal use only, etc.) *Specification of this descriptor is quite straightforward.*
12 Date	Input year of publication of data-source *This descriptor refers to the publishing date of the data-source, not the actual CRM data, since a separate descriptor is created for that purpose, see 'years covered' descriptor.*
13 Format	Select from term-list (PDF/WORD/EXCEL/ACCESS/OTHER) *Specification of this descriptor is quite straightforward.*
14 Identifier (DOI/ISBN/URL)	Input/Copy-paste *Specification of this descriptor is quite straightforward. In Endnote, there are three fields for the Identifier descriptor: (DOI/ISBN/URL)*
15 Language	Select from term-list European Language + Jap & CHN *For this descriptor, a short-list of languages is created, being the EU languages expanded with Japanese and Chinese. If another language were to be encountered, manual input is required.*

Table 2 - The 15 Dublin Core metadata fields and their definition

Finally, the documents that are part of the Knowledge Base can be entered in two ways. The first one was developed when starting to build the portal and allows the user to enter manually the description (metadata) of the document, using a form presenting the 15 fields of Dublin Core (Figure 17). The second one uses XML files to import by batch the documents. It was modified to accept EndNote exports. These exports allow the partners of the project to import by batch the documents they have already described in the EndNote application.

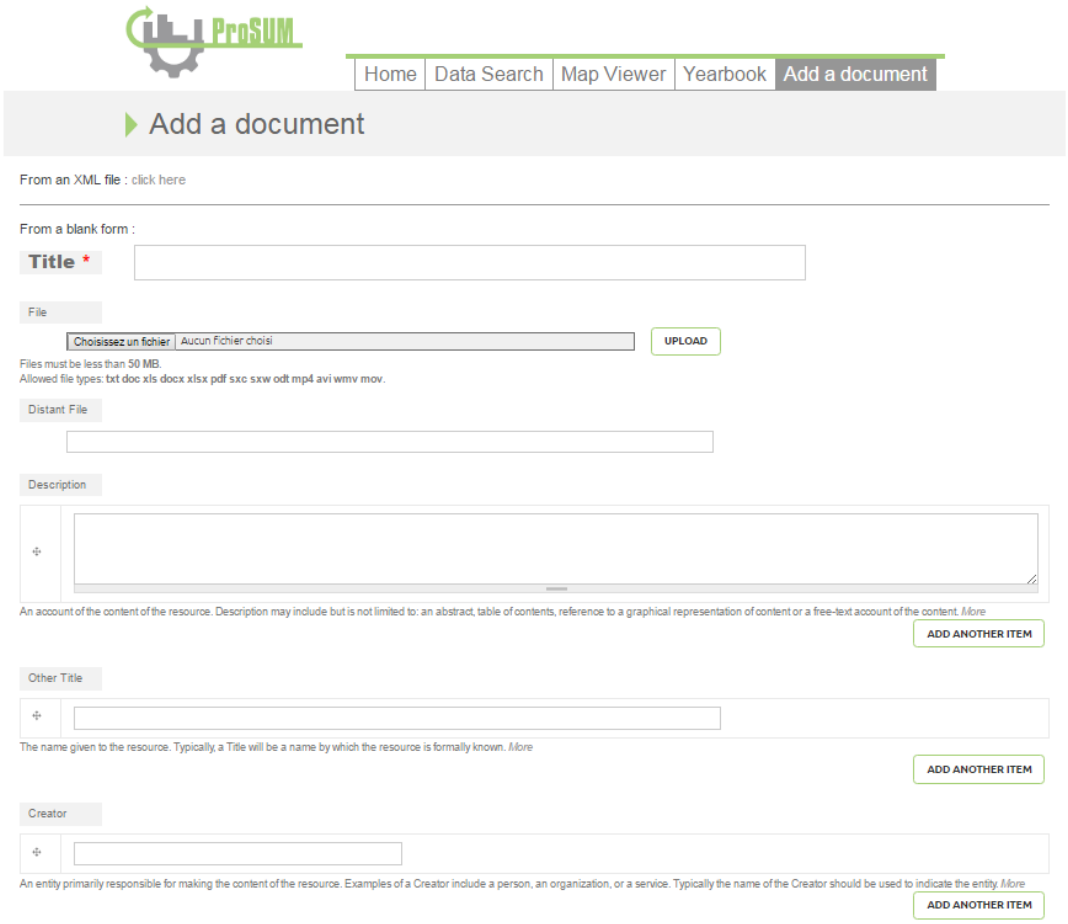


Figure 17 - Add a document to the Knowledge Base

3.7 – The Metadata Catalogue of the structured data

The ProSUM Metadata Catalogue is the central access point to metadata concerning data on secondary sources of raw materials from "urban mines" - data on WEEE, ELV, batteries and mining waste (see Figure 18Figure 18). It provides tools for compilation of those metadata in a standardized format that allows users to effectively search through the database.

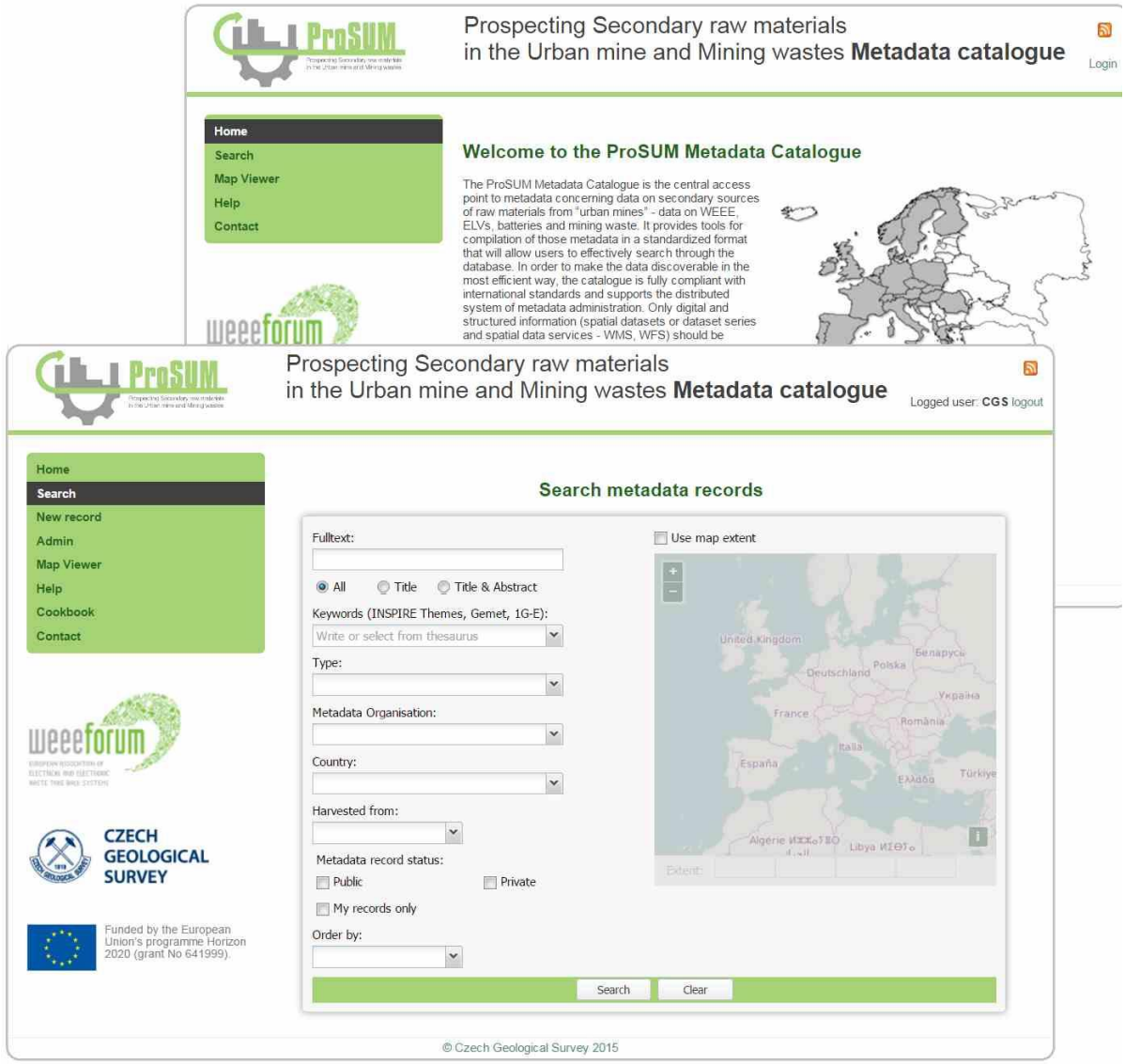


Figure 18 - The updated version (mid-2016) of the Metadata Catalogue layout

In order to make the data discoverable in the most efficient way, the catalogue is fully compliant with international standards (ISO 19115 Geographic Information: Metadata, ISO 19119 Geographic Information: Services, ISO 19110 Geographic Information: Methodology for feature cataloguing, ISO 15836:2009 - Information and documentation - The Dublin Core metadata element set) and supports the distributed system of metadata administration. The profile is compliant with the INSPIRE Metadata Implementing rules: Technical Guidelines from 2013-10-29.

The advantage of the distributed system is the possibility to reuse already existing metadata from the European, national or related projects catalogues (e.g., Minerals4EU, One Geology, etc.). This implies less additional work for the data providers as they should already have provided INSPIRE compliant metadata of their data. They would only have to add English translations (if not existent) and project-specific keywords. The possibilities on how metadata can be inserted and maintained in the catalogue are presented in Figure 19.

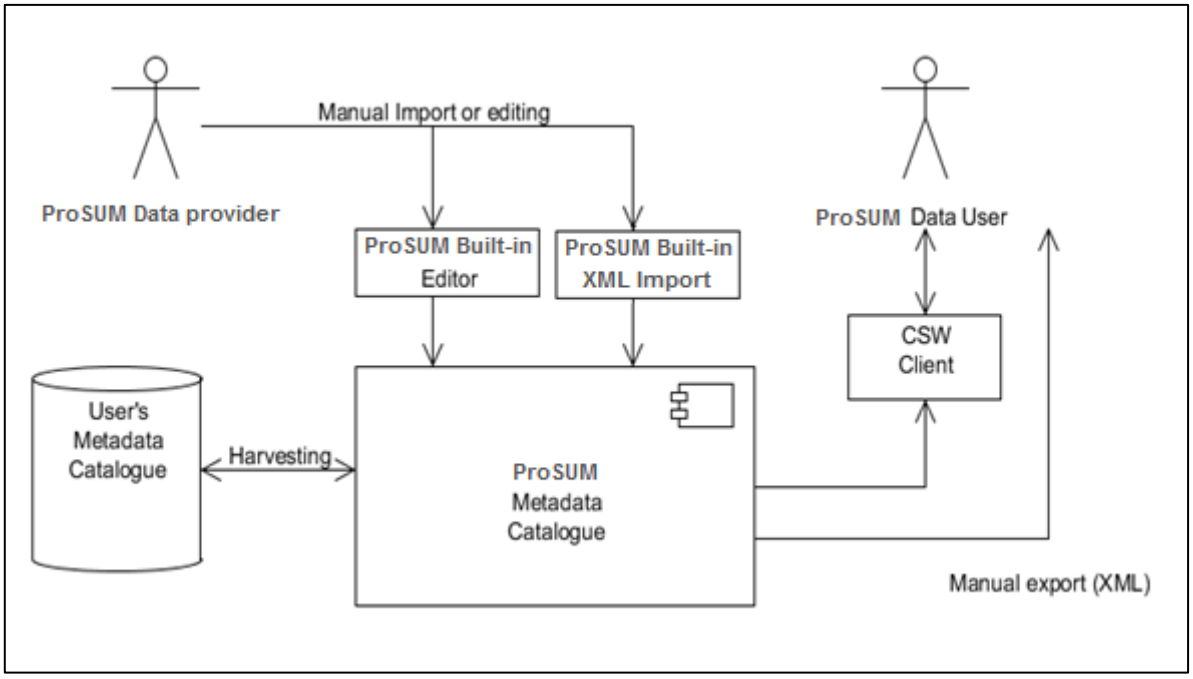


Figure 19 - The strategy for the ProSUM metadata catalogue architecture

Only digital and structured information (spatial datasets or dataset series and spatial data services - WMS, WFS) should be described by metadata in this catalogue.

Based on experience from the previous projects (Minerals4EU, OneGeology-Europe, etc.) implementation of the following components of metadata catalogue is recommended:

- searching tools (search by keywords, data source type, organization, full text and country);
- editing tools for creating, updating and maintaining the metadata records;
- administrative tools for harvesting, management metadata profile and access rights;
- help, cookbook;
- Metadata keywords code list – structure, using rules of semantic web (URI), linking with M4EU codes and EU legislation and directives should be discussed;
- Catalogue Service for Web (CS-W 2.0.2.) - service for online integration of the metadata records with the other portals (e.g., Minerals4EU, OneGeology-Europe, INSPIRE or other national and project-specific geoportals);
- Support and validation of INSPIRE metadata profile.

In order to display a metadata record for which an on-line map service is available, the Metadata Catalogue is integrated into the ProSUM Portal. The catalogue enables systematic discovery, viewing and use of data on WEEE, ELV, BATT and mining waste across Europe.

3.8 – The search facilities

The search facilities are based on a search engine indexing the Central Diffusion Database, the documents corpus (non-structured data) and some external databases (if any). A user interface allows the end users retrieving data from the whole EU-UMKDP using a simple input (Google-like search, a simple sentence will search in the whole EU-UMKDP), or using some specialized interfaces (search specific concepts using their main attributes). Almost all the indexed concepts have a geographic and temporal extents (coming from INSPIRE MR/ERML and other data models for the Diffusion Database, coming from Dublin Core metadata for the documents). These extents allow the user retrieving the most accurate response for his search.

3.9 – EU-UMKDP dedicated services

In order to answer the end users most common requests, some specific calculation services have been set up for ProSUM on top of the Central Diffusion Database. Then, some applications have been developed to allow users to easily find the best information.

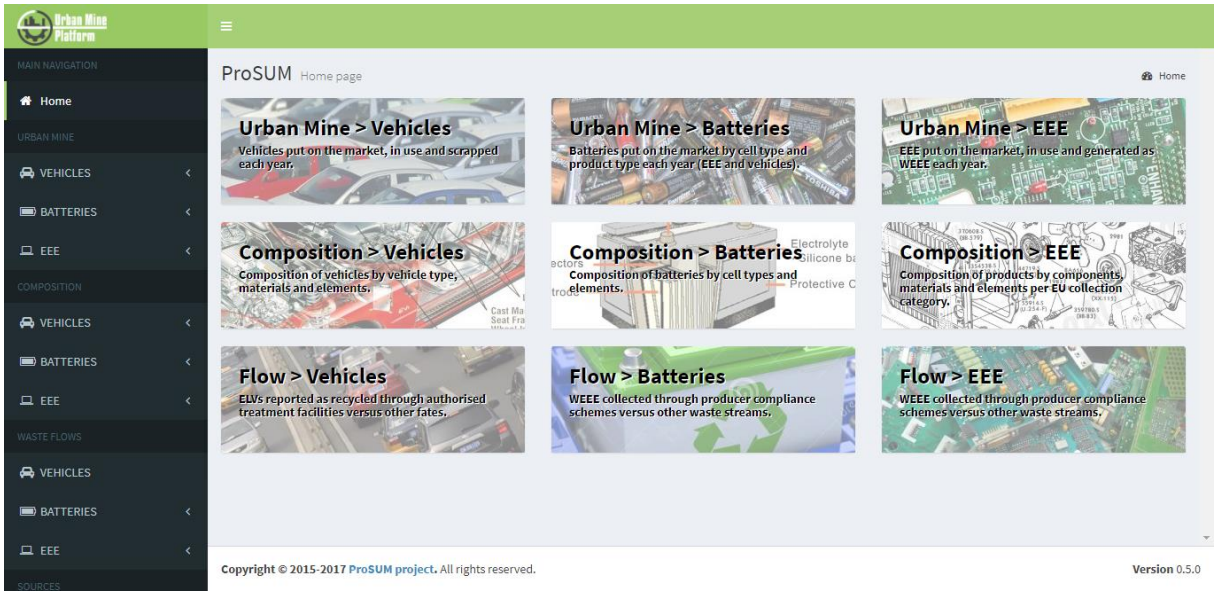


Figure 20 - Homepage of ProSUM applications (http://www.urbanmineplatform.eu)

The homepage of the ProSUM applications section presents in 9 tiles the access to the different applications (

Figure 20). These applications are split in three topics:

- the '**Urban Mine**' gives information about the products Put on the Market, the content of the Stocks and the Wastes generated;
- the '**Composition**' describes the composition of the products in components, materials, elements;
- the '**Waste Flow**' displays the values of the different flows and the potential gaps.

These topics are then subdivided to present the information of the three waste groups (Vehicles, Batteries, EEE).

Depending of the waste group, the applications proposed are not always the same: some waste groups have some particularities that request some specific application(s) or that cannot be displayed in any application. For example, the composition of batteries expressed in terms of components or materials is irrelevant and such related application has not been integrated in the system.

3.9.1 – Urban Mine

Urban Mine applications present figures related to the Put on the Market, Stock and Waste generated.

Depending of the waste group selected, the values can be aggregated in different ways:

- Vehicles: the products are aggregated using the power train type (Diesel, Petrol, BEV/Fuelcell, HEV, PHEV, LPG, NG…);
- Batteries: the products are aggregated using the battery type (Primary Lithium based, Rechargeable lithium based, Nickel-Cadmium…) or for more detail, when a battery type is selected (e.g., Rechargeable lithium based), the combination of composition and application can be selected (e.g., Lithium cobalt dioxide (LiCoO2) Laptops, Lithium cobalt dioxide (LiCoO2) Mobile phones, Lithium nickel manganese cobalt (LiNMC) PHEV…);
- EEE: the products are aggregated using the collection category (Temperature exchange equipment, Screens, Lamps, Large equipment, Small equipment, Small IT).

Three representations of the figures are proposed, the first one presents the data in pieces of products, the second one in weight and the last one in weight per capita. These representations are really complementary as for example for EEE, in the representation in pieces the Lamps can depict a huge value compared to the Temperature exchange equipment but the ratio is inversed when we look at the representation in weight. Also the representation in weight per capita allows easy comparison of the European countries independtly of the size of the country.

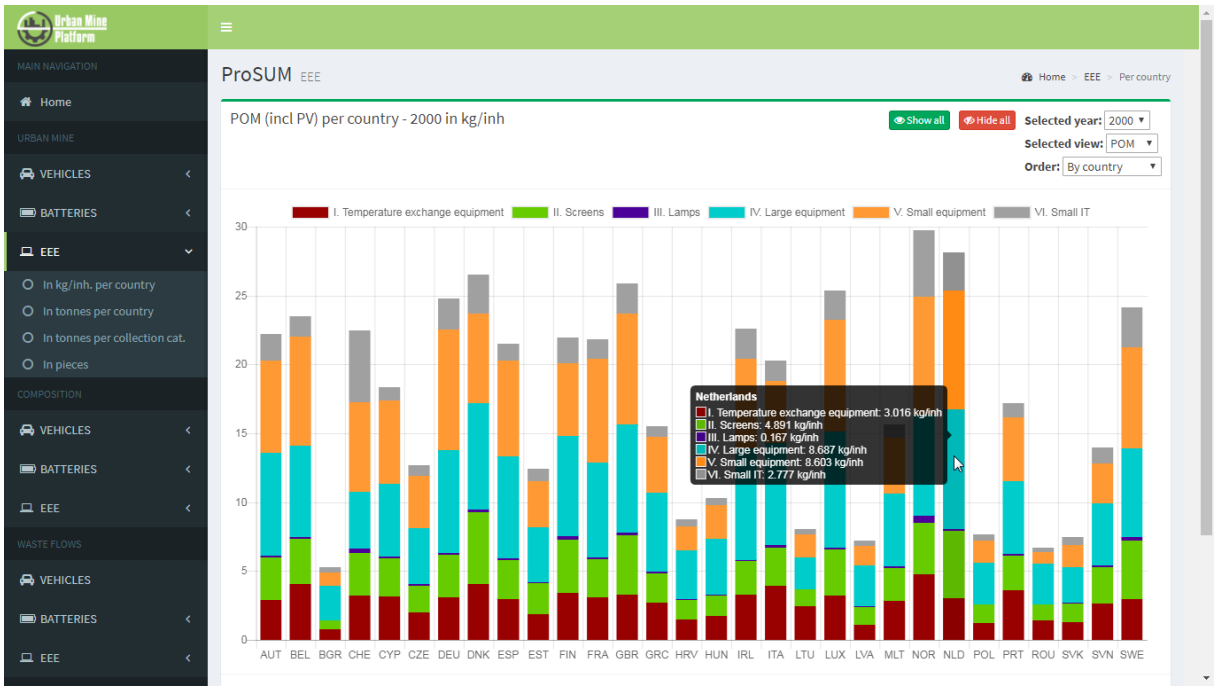


Figure 21 - ProSUM applications - Urban Mine for EEE in kg/capita (sorted by countries)

The graphs propose some functionalities to help the user to retrieve easily the information. Figures are displayed over on the graph in tooltips (
Figure 21). The different data series of the graph can be shown or hidden just with a click on the item in the legend. Two buttons at the top of the graph allow also to directly show or hide all series (to hide all countries in the country applications to then select only one or two to compare them). For applications presenting figures by country, an option allows also to sort the values to easily order the countries from the one with the highest value to the one with the lowest (or reverse) (

Figure 22).

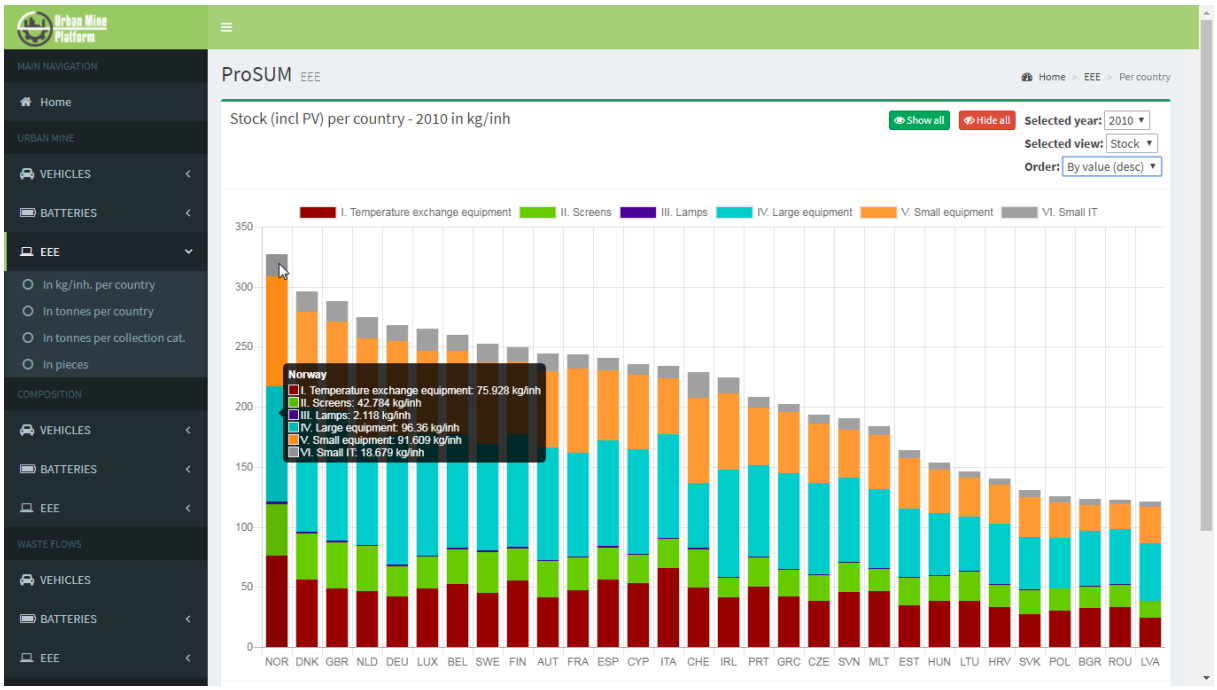


Figure 22 - ProSUM Applications - Urban Mine for EEE in kg/capita (sorted by valies)

The options of the graph can be changed using the select box on the top right of the application. These options propose to change the view (from Put on the Market to Waste passing by Stock). Depending of the application and the waste stream, these options can propose to change the country visualized, the collection category or the type of batteries.

Depending of the applications, the chart can be presented as a bar chart, presenting factual values for the period or as a line chart, presenting some values that can be projected (

Figure 23). Also, a grey background indicates if the values presented on it are projected or not, and an indication in the explanation area clarifies the meaning of the symbols or colours used in the application.

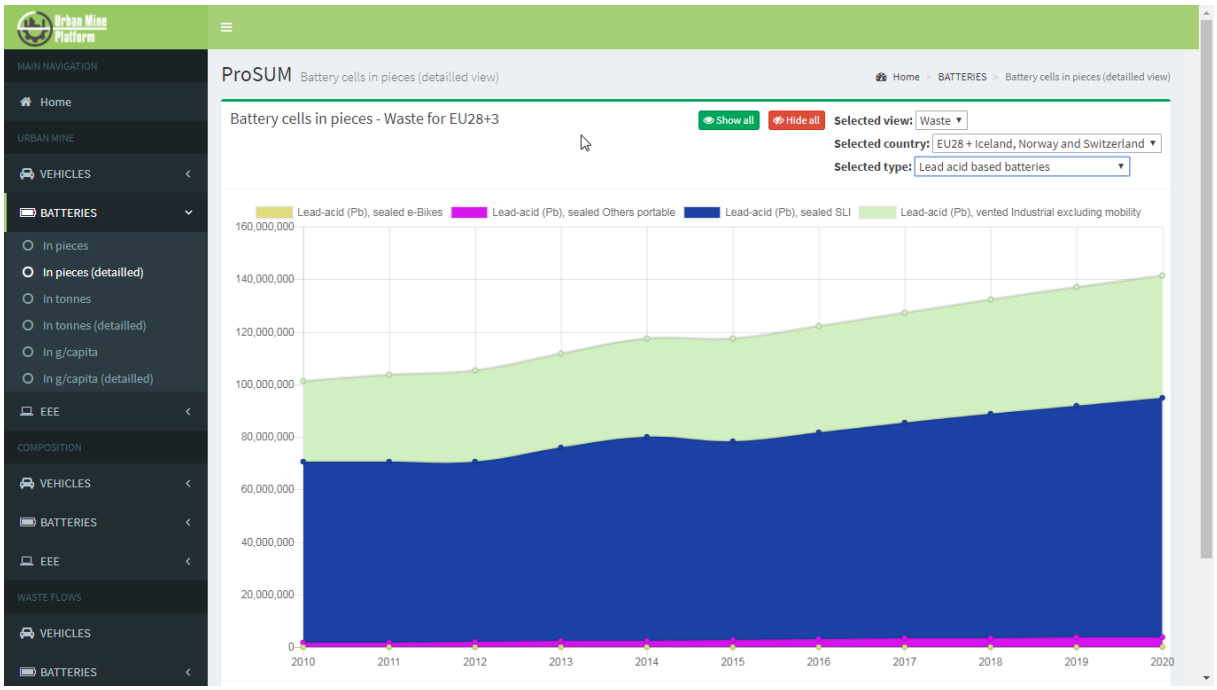


Figure 23 - ProSUM Applications – Urban Mine for Batteries in pieces detailed view for Wastes generated

3.9.2 – Composition

Composition applications present, for the selected parameters, the composition in components, materials or elements for the products Put on the Market, in the Stocks or in what will become Wastes generated for the specified year (

Figure 24).

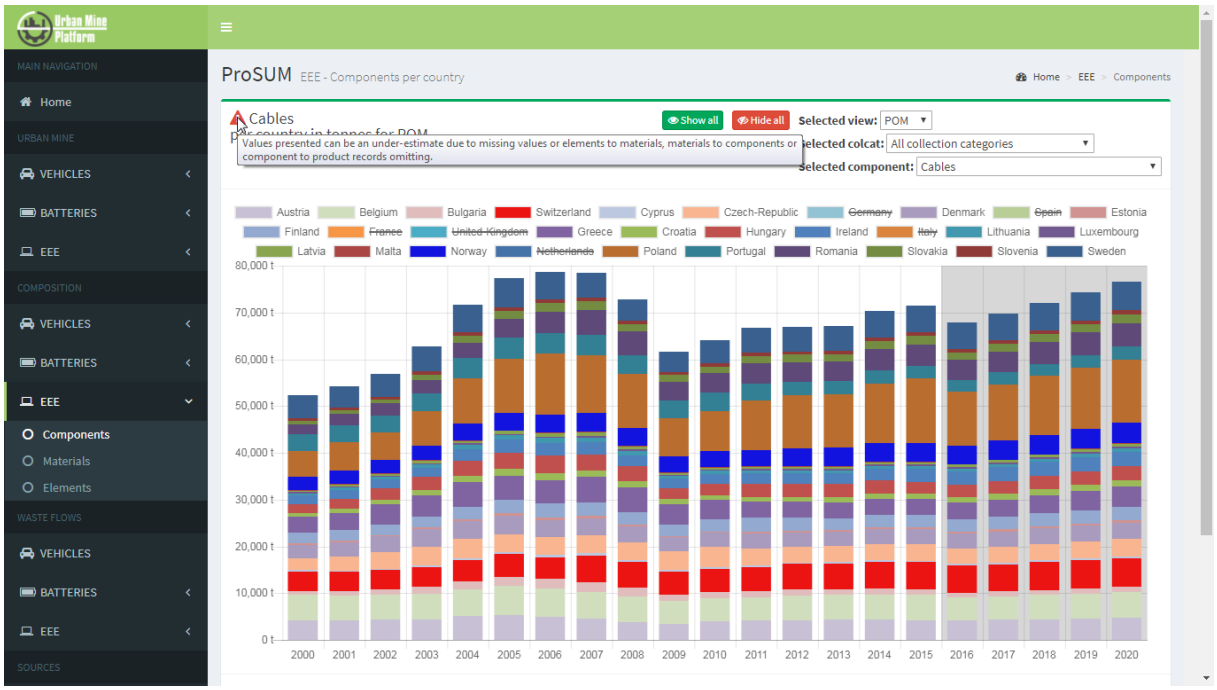


Figure 24 - ProSUM Applications - Composition for EEE in components for Car Cables

Vehicles and EEE waste groups propose the composition in components, materials and elements when Batteries only propose the composition in elements.

The user can select the composition/material/element he wants to visualize on the graph and depending of the waste streams, restrict the visualization to some specific products:

- For Vehicles: the filter is done on the energy type of the motor;
- For Batteries: the filter can be done at three levels. On the type of battery, then on the composition of the battery and finally on the application of the battery (
- Figure 25);
- For EEE: the filter is made on the collection category.

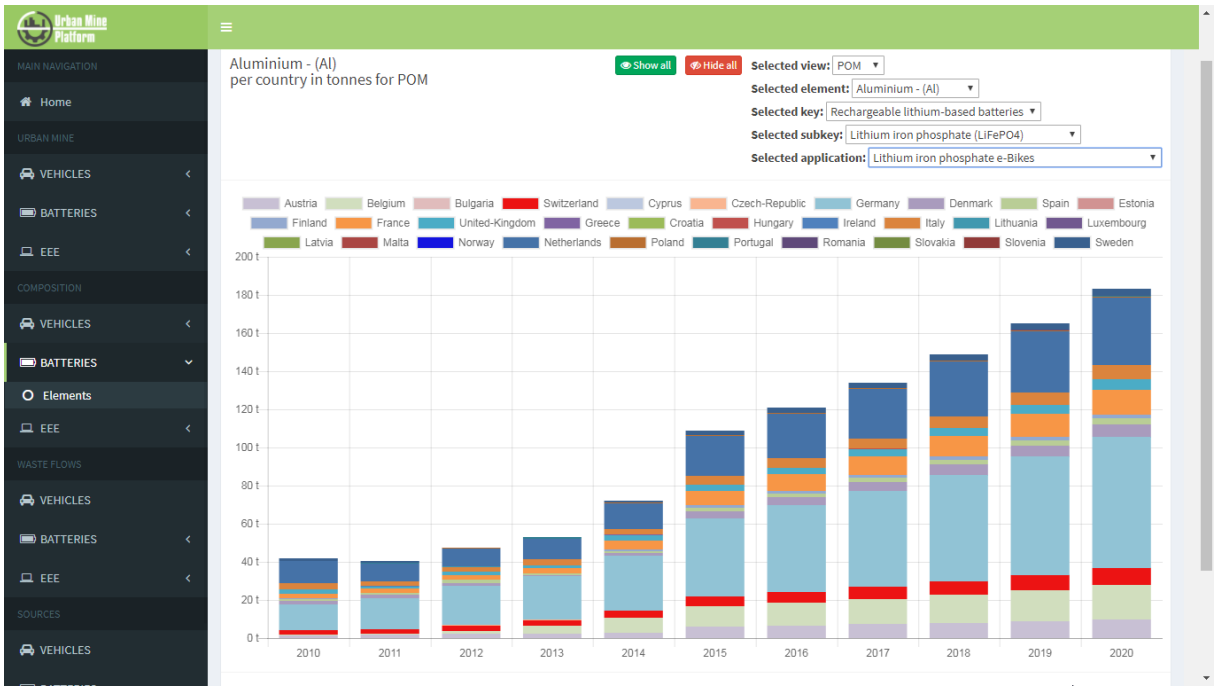


Figure 25 - ProSUM Applications - Composition for Batteries in Elements

3.9.3 – Waste flows

Waste flows applications present the flows of products after they become waste and allow visualizing the different types of flows and the gaps between the generated wastes and the really known flows of waste (
Figure 26).

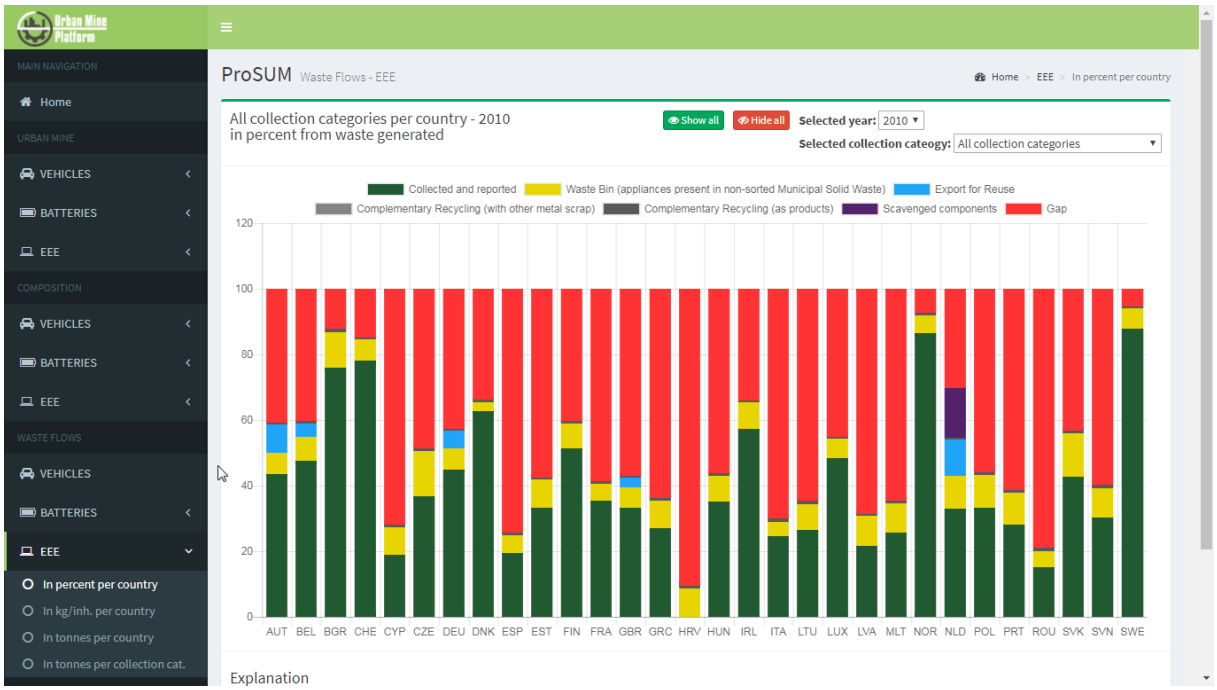


Figure 26 - ProSUM Applications - Urban flows for EEE, view in percent per country

Other visualization modes allow the user to view for a selected flow the difference between the countries, the years or the type of product (energy motor type, batteries type, collection category). These visualizations allow notably to view easily the evolution of the flows in time or to compare the countries between them (

Figure 27).

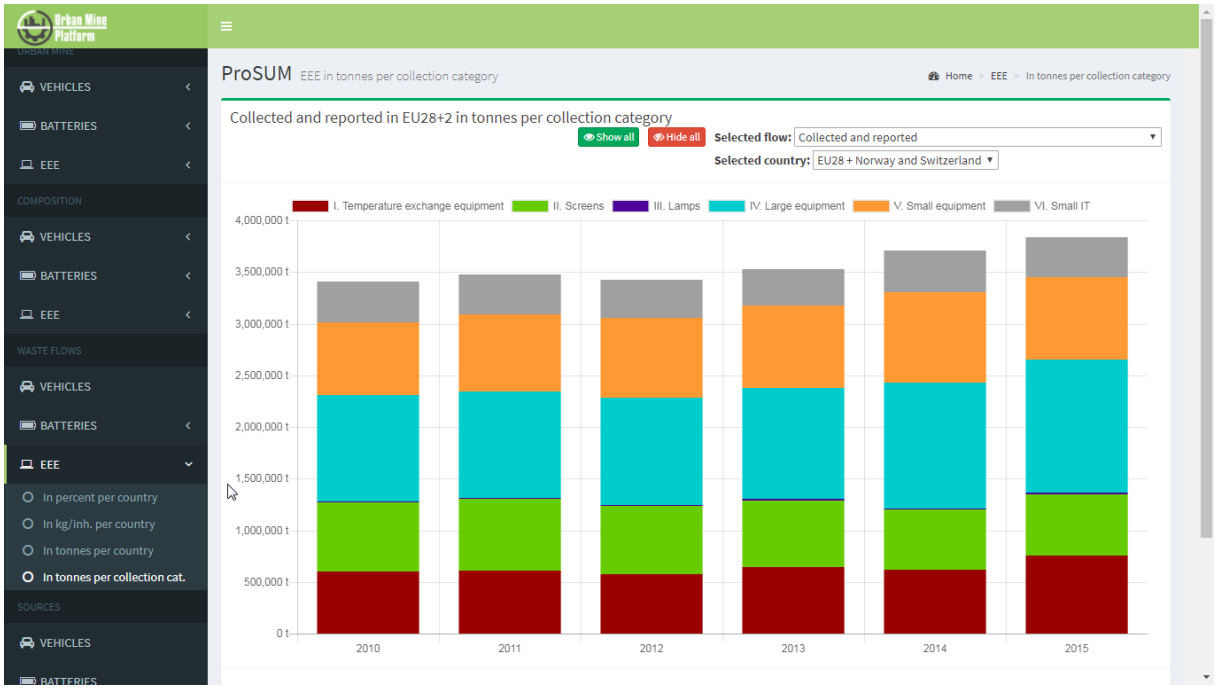


Figure 27 - ProSUM Applications - Waste flows for EEE, view in tonns per collection categories

All the applications are accompanied by the possibility to view more information about the data they are presenting via a dedicated button in the Explanation section. This button leads to related records of the Metadata Catalogue for structured data.

4 – The EU-UMKDP and the MICA Expert System (the EU-RMICP)

The MICA Expert System is currently being developed in the frame of the H2020 MICA project and more precisely within its work package 6 entitled 'The European Raw Materials Intelligence Capacity Platform (EU-RMICP)' whose partners are BRGM, BGS, GeoZS, GEUS, GTK, JRC and LIG (Laboratoire d'Informatique de Grenoble).

The description of the project can be found on the project website (http://www.mica-project.eu/). To briefly summarize: MICA (Mineral Intelligence Capacity Analysis - 2016-2017) has among its objectives to develop a platform of knowledge, the EU-Raw Materials Intelligence Capacity Platform (or EU-RMICP), integrating metadata on data sources related to primary and secondary mineral resources and bringing the end users with an expertise on the methods and tools used in mineral intelligence. In practice, the system owes be capable of bringing relevant 'answers' of the type 'how to proceed for...' on almost any question relative to mineral resources, on the whole supply chain, since the prospecting until the recycling, taking into account the environmental, political and social dimensions.

To meet this challenge, the EU-RMICP is based on an ontology of the domain of mineral resources (coupled with more generic cross-functional ontologies, relative to commodities, time and space), which represents the domain of the questions of the users (experts and non-experts). The user navigates this ontology by using a Dynamic Graph of Decision (DDG) which allows him/her to discover the solutions which he/she is looking for without having to formulate any question. The system is coupled with a 'RDF TripleStore', a database storing the ontologies, factSheets, docSheets and flowSheets (i.e., specific formatted forms) related to methods and documentation, scenarios and metadata (Figure 28Figure 29 The MICA Expert System. General architecture of the EU-Raw Materials Intelligence Capacity Platform.).

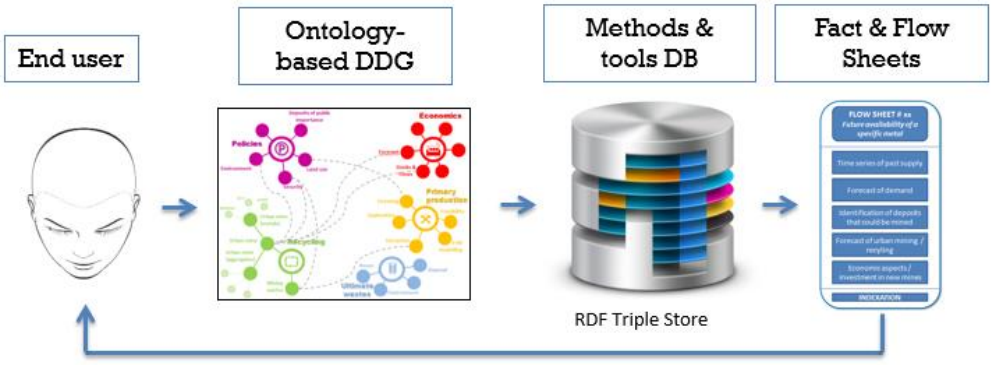


Figure 28 - The MICA Expert System, and the ontology-based Dynamic Decision Graph (DDG).

This particularly innovative system can be widened (perimeter or scope and granularity) and represents a prototype of a modern expert system.

In practice, this system will be connected with the existing Knowledge Data Platforms, e.g., the IKMS (EURARE), the EU-MKDP (Minerals4EU), **the EU-UMKDP (ProSUM)**, the EU-CRMKDP (SCRREEN), the European Geological Data Infrastructure (EGDI) developed by EuroGeoSurveys (EGS) and the RMIS 2.0 (Raw Materials Information system) currently being developed by the European Commission DG JRC in Ispra, allowing them to make their users benefitting of the Expert System (see Figure 29 The MICA Expert System. General architecture of the EU-Raw Materials Intelligence Capacity Platform.Figure 29).

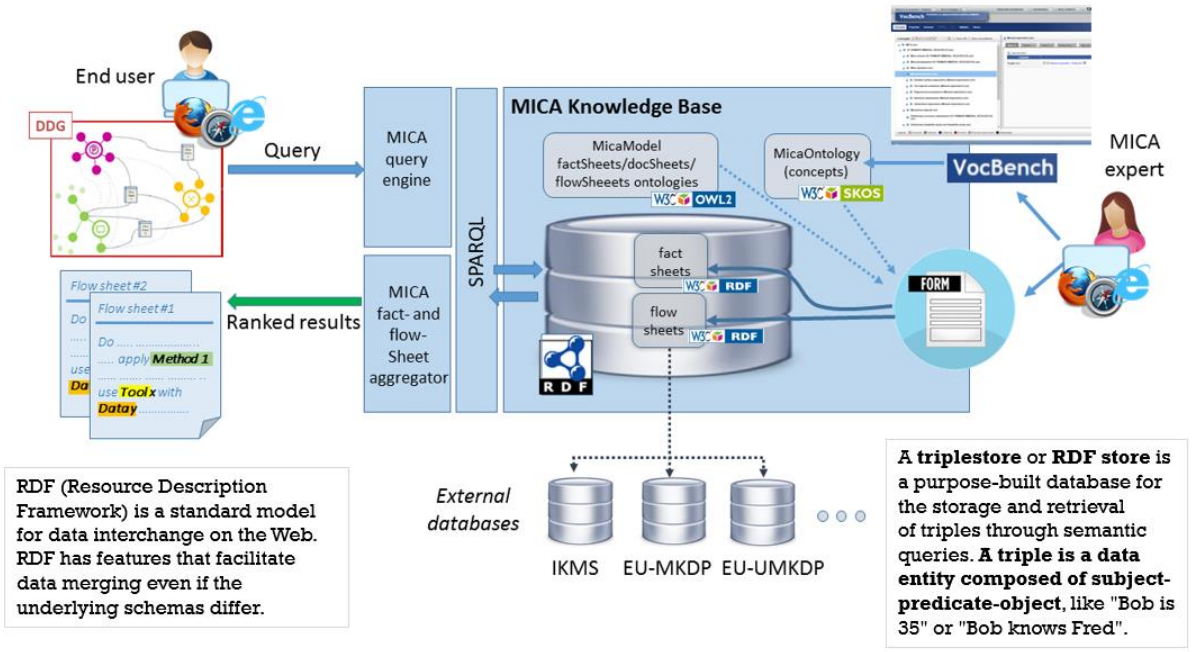


Figure 29 The MICA Expert System. General architecture of the EU-Raw Materials Intelligence Capacity Platform.

The MICA Main Ontology actually covers 7 thematic domains: 'Primary' and 'Secondary Mineral Resources', 'Industrial Processing and Transformation', 'Raw Materials economics' (including CRMs), 'Raw materials Policy & Legal Framework', 'Sustainability of Raw Materials' and 'International Reporting' (Figure 30). The DDG offers in a single place, a unique access to most of the data available, including a contextual access to resources like the European legislation, and an

access to several key studies like the Minventory study[4] (Parker et al., 2015), or the Material System Analysis[5] (BIO by Deloitte, 2015), etc.

One should note that the first functionality of the DDG is not to be a 'pure' search engine. However, data, information and knowledge being strongly and cleverly connected, the DDG and its side applications can play the role of an 'intelligent' search engine in the Raw Materials domain, and also act as a powerful decision-aid tool.

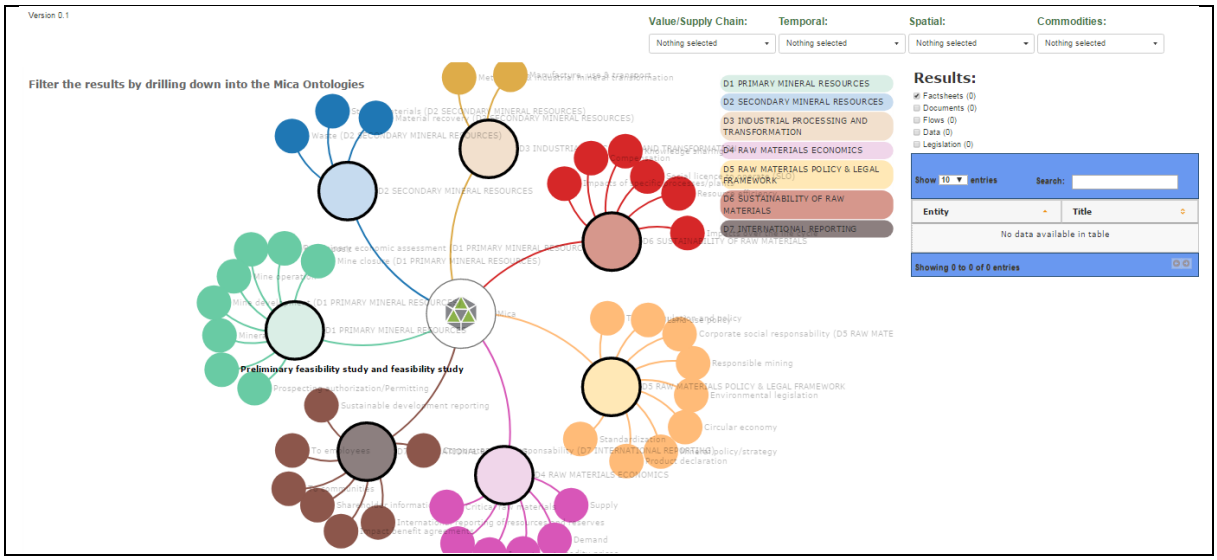


Figure 30 The DDG interface showing the Main Ontology and the 7 domains covered, representing about 300 conceptsand sub-concepts.

Figure 30 shows the DDG interface with the Main Ontology and the 7 domains covered, representing about 300 concepts and sub-concepts. Transversal ontologies, i.e., 'Value_Supply Chain', 'Temporal', 'Spatial' and 'Commodities' are used as filters allowing to speed up the process of retrieval of doc/fact/flowSheets and linkSheets. Results are presented on the right side and are ranked by pertinence.

FactSheets are used to describe in detail a method or a tool, giving all the necessary information to the end user on how to implement this method for resolving a problem. DocSheets can be seen as a complementary source of information, explaining some concepts (which are not methods or tools) such as substitution, criticality… an end user may appreciate to find during his navigation on the Dynamic Decision Graph. FlowSheets can be seen as 'cooking recipes' allowing to answer complex queries an end user may have and which necessitates to link in a certain order several factSheets and related data.

FactSheets and flowSheets indicate what type of data is necessary for running the method(s) and what are their source(s). An exhaustive inventory of data/data sources has been realized, which

[4] http://ec.europa.eu/growth/tools-databases/newsroom/cf/itemdetail.cfm?item_id=8273&lang=en
[5] https://ec.europa.eu/jrc/en/scientific-tool/msa

provides the system with detailed metadata related to these data sources. Each metadata is internally indexed (i) to the fact/doc/flowSheet using this source and (ii) to the domain(s)/concept(s)/sub-concept(s) to which it may be useful.

In practice, the collaboration between the EU-UMKDP and the EU-RMICP can work in two complementary ways, **for the mutual benefit of the two platforms**:

- The EU-UMKDP will appear as a referenced data source for the urban mine in the DDG (see Figure 29);
- The EU-RMICP could improve its capacity to answer regarding the urban mine if it is feeded with dedicated fact and docSheets prepared by the ProSUM team and if the perimeter and the granularity of its Main Ontology are augmented.

As fact and docSheets are annotated with concepts, other sheets, references, metadata related to data sources, the EU legislation, etc., contributing to the EU-RMICP would also contribute to make the EU-UMKDP the **REFERENCE SOURCE** for the Urban Mine.

5 –CONCLUSIONS & PERSPECTIVES

The European Union - Urban Mining Knowledge Data Platform, the EU-UMKDP represents the first attempt to create a one-stop shop or a one-stop information gateway to urban mine data and knowledge. The system is actually conceived to manage both structured data (i.e., from databases) and also semi- or non-structured data (i.e., all types of documents from reports, monographs, graphs and bar charts, and images to videos, nearly whathever the format).

However, its main asset will remain invisible for most of the users and lies in the data model which has been developed for data management and database development. This data model is quite detailed, allowing to manage at the same time products and compositions, and stocks and flows, and generic enough to manage concomitantly WEEE, ELV and BATT. This also means that this data model has the intrinsic characteristics for being extended in the future to other types of urban wastes.

The EU-UMKDP is not an isolated Knowledge Data Platform. It is connected to the EU-MKDP, the platform developed in the Minerals 4EU project and which provides the ProSUM KDP with data related to primary mineral resources and mining wastes, making thus the EU-UMKDP one of the most powerful hubs in terms of secondary resources observatory and CRM potential evaluation in Europe,

The functionality of the portal can even be extended by coupling it to the MICA Expert System. This coupling allows a user to learn how to proceed to get relevant data and information related to primary and secondary material flows. The two platforms are linked by developing dedicated doc/fact/flowSheets related to the various aspects of the urban mine for the MICA EU-RMICP. This in turn will contribute to making the EU-UMKDP an indispensable ‘player’.

REFERENCES

1. Cassard D., Tertre F., Schjøth F., Heijboer T., Čápová D., Hallberg A., Šinigoj J. (2016). The ProSUM EU-Urban Mining Knowledge Data Platform Specification (T5.3.1, part of D5.7). ProSUM Public Report, 194 p.

2. Heijboer T., Schjøth F., Podboj M., Cassard D., Tertre F., Hallberg A. (2017). Deliverable D5.5: Data models and code lists. ProSUM Public Report, 1157 p.

3. INSPIRE Directive (2007). http://inspire.jrc.ec.europa.eu/; http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF.

4. Serrano J.-J., Kauniskangas E., Cassard D. (2010). ProMine: architecture, portal and web services to provide a European mineral resources information system, 4th INSPIRE Conference. Krakow, 22-25 June.

5. Cassard D., Bertrand G., Billa M., Serrano J.-J., Tourlière B., Angel, J.-M., Gaál G. † (2015). ProMine mineral databases: new tools to assess primary and secondary mineral resources in Europe. P. Weihed (ed.), 3D, 4D and Predictive Modelling of Major Mineral Belts in Europe, Mineral Resource Reviews, DOI 10.1007/978-3-319-17428-0_2, pp. 9-58.

6. Gruijters S. and Fruijtier S. (2011). Using cloud computing to increase the usability of distributed Web GIS Systems. eChallenges e-2011 Conference Proceedings, 26-28 October 2011, Florence, Italy. Paul Cunningham & Miriam Cunningham (Eds.), IIMC International Information Management Corporation, 2011, ISBN: 978-1-905824-27-4, 8 p.

7. EuroGeoSource (2013). EuroGeoSource homepage http://eurogeosource.eu/. In the menu EuroGeoSource Deliverables: 1) Publications and presentations: http://eurogeosource.eu/deliverables/publications-and-presentations; 2) Public Reports: http://eurogeosource.eu/deliverables/public-reports.

8. INSPIRE Thematic Working Group Mineral Resources (2013). D2.8.III.21 Data Specification on Mineral Resources – Technical Guidelines. 156 p. European Commission Joint research Center Publisher. http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_MR_v3.0.pdf.

9. EarthResourceML (2013). http://www.earthresourceml.org/earthresourceml/2.0/doc/ERML_HTML_Documentation/

10. Vuollo J., Cassard D., Simons B., Seymon A. (2010). The Earth resource data exchange model (EarthResourceML) – a tool for delivering INSPIRE and ProMine mineral resource data, 4th INSPIRE Conference. Krakow, 22-25 June.

11. Vuollo J., Simons B., Laxton J., Cassard D., Seymon A. (2012). EarthResourceML v.2.0 – an upgrade of the CGI-IUGS Earth Resource Data Model due to INSPIRE Data Specification. 34th International Geological Congress, Brisbane, August 2012, Abstracts on CD-ROM.

12. Cassard D., Serrano J.-J., Vuollo J. (2012). The use of the standard exchange EarthResourceML in the ProMine project. 34th International Geological Congress, Brisbane, August 2012, Abstracts on CD-ROM.

13. Raymond O., Vuollo J., Richard S. and CGI Interoperability Working Group (2012). Global Geoscience Data Transfer Standards - GeoSciML and EarthResourceML. eResearch Australasia Conference, Brisbane, Australia.
http://conference.eresearch.edu.au/eres2012/posters-2012/

14. Straalen van V., Huisman J., Habib H., Chancerel P., Maehlitz P., Rotter S., Wäger P., Schjøth F., Hallberg A., Scheepens A. and Cassard D. (2015). Review and harmonisation of data. Deliverable 5.3.. ProSUM Public Project Report, 97 p.

15. Schubert C., Vuollo J., Tomas R., Cassard D. and WP5 Partners (2014). Minerals Intelligence Network for Europe – Minerals4EU-WP5. WP5: Common terminology for Minerals4EU – Version 1.0, Minerals4EU Public Report, 330 p.