MINERAL INTELLIGENCE
FOR EUROPE

Deliverable D5.3.1

Specification of steps needed for the integration of the e-Minerals Yearbook in the Minerals4EU database

Authors and affiliation:
Jasna Šinigoj (GeoZS), **Daniel Cassard** (BRGM), **Mikael Pedersen** (GEUS)

E-mail of lead author:
jasna.sinigoj@geo-zs.si

Version: 23-01-2019

This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation programme under grant agreement number 731166.

Deliverable Data		
Deliverable number	D5.3.1	
Dissemination level	Public	
Deliverable name	Specification of steps needed for the integration og the e-Minerals Yearbook in the Minerals4EU Database	
Work package	WP5, Improvement of KDPs' applications and interaction with the RMIS and the GeoERA Information Platform	
Lead WP/Deliverable beneficiary	GeoZS	
Deliverable status		
Submitted (Authors)	23/01/2019	Jasna Šinigoj, Daniel Cassard, Mikael Pedersen
Verified (WP leader)	30/01/2019	Mikael Pedersen, Daniel Cassard
Approved (Coordinator)	30/01/2019	Jørgen Tulstrup

GENERAL INTRODUCTION

This report describes the steps to be undertaken to integrate the e-Minerals Yearbook (e-MYB) into the Minerals4EU database (EGDI database) and assesses the possibilities and technical solutions related to setting up harvesting systems and automation routines. This report is based on a draft of the H2020 ORAMA deliverable D3.1 INSPIRE "Compliance of improved datasets".

EXECUTIVE REPORT SUMMARY

One of the main goals of the Mintell4EU is to integrate the e-MYB with the overall data infrastructure of the EGDI, which forms the basis for GeoERA Information Platform. Having the e-MYB maintained in the EGDI database, that already holds the Minerals Inventory will ensure its long-term sustainability. The data from the e-MYB and Minerals Inventory will be disseminated through a number of channels including the EGDI portal and the already existing Minerals4EU portal. (

Figure 2).

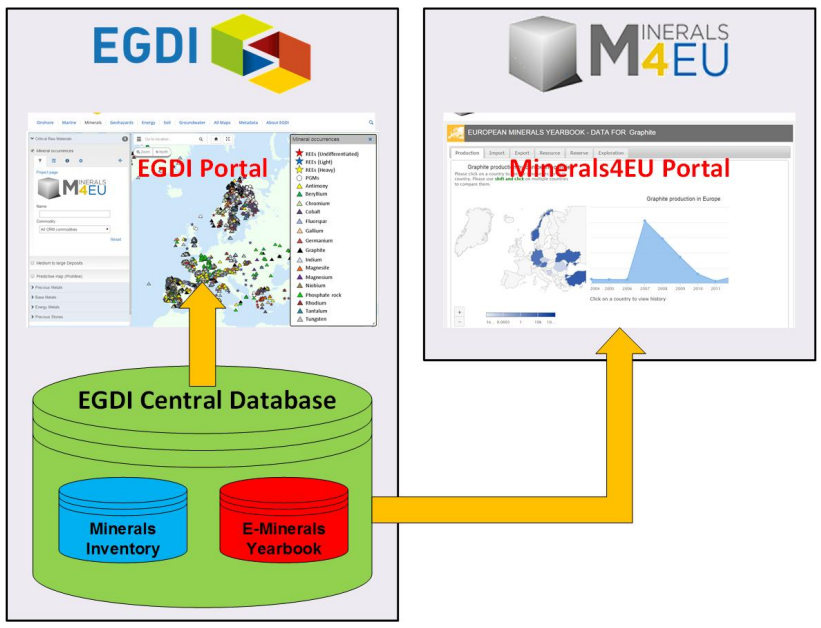

Figure 1. *The end goal of Mintell4EU is to integrate the E-MYB with the Minerals Inventory in the central EGDI database from which data are disseminated through the EGDI portal, the minerals4EU portal and other channels*.

A new e-Minerals Yearbook (e-MYB) data model (DM) will be created that will reuse parts of the BGS website data model for production and trade (import and export data). This DM will thus be jointly developed by BGS, BRGM, GeoZS and GEUS. It will be used (i) to harvest National Providers databases (DBs), (ii) to build a Web application for countries/institutions not having harvestable data from DBs and (iii) to build the Harvesting database. BGS will receive a dump of this DB which will be used in the following way: (i) the creation of a copy of the GeoZS' Harvesting DB to store exploration, reserves and resources data, (ii) extraction from this DB of production data (if harvested/provided) which will then be incorporated into the BGS World Mineral Statistics DB. Validated data of the 5 types of data, coming thus from two different DBs (exploration, resources & reserves data, BUT NOT PRODUCTION, from the BGS Copy of the Harvesting DB, and production and trade data from the BGS website DB), will be provided by BGS to BRGM for the update of the previous version of the e-MYB via a

Web service based on the Minerals4EU/ORAMA e-Minerals Yearbook data model (Figure 2).

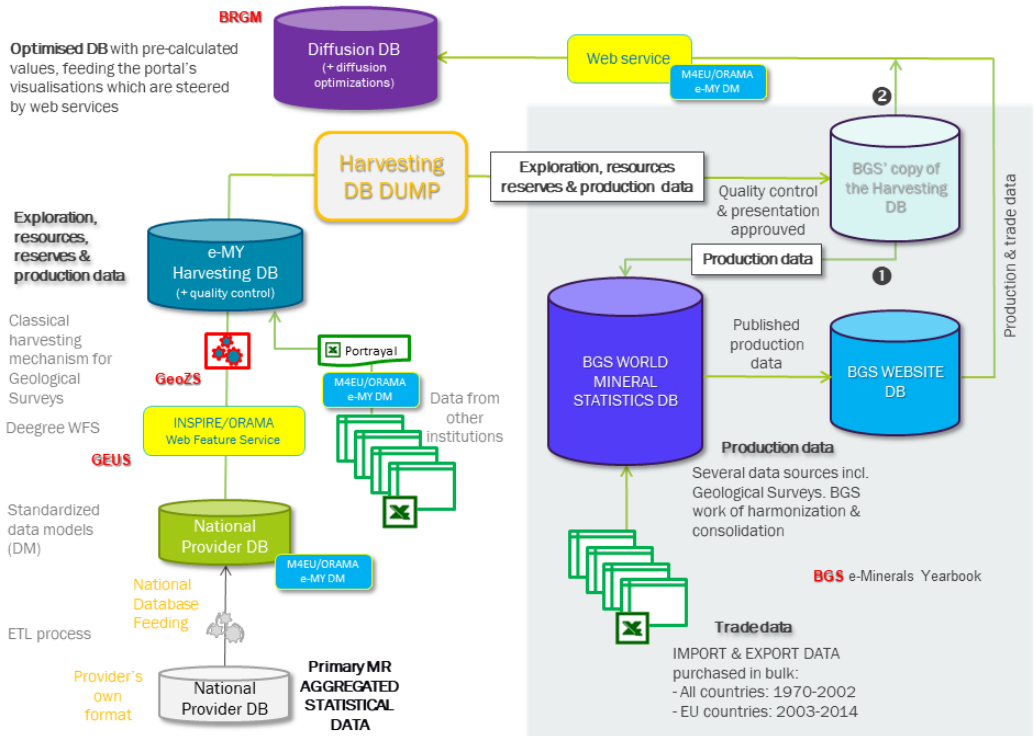

Figure 2. *The e-Minerals Yearbook feeding data flow (after Cassard et al., in prep.).*

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1 OVERVIEW

The integration of the e-MYB with the overall GeoERA Information Platform (the EGDI database) is one of the main goals of the Mintell4EU project and a key activity in WP5. Recommendations for this integration have been drafted in the ORAMA project and are described in more details in ORAMA deliverable D3.1.

The INSPIRE data model contain no elements for aggregated statistical data about mineral resources that is required for the Minerals Yearbook hence two possibilities were exploited by the ORAMA project:

- The ERML Lite approach;
- The O&M approach.

The ORAMA recommendation (established by a group of Experts during the first ORAMA Progress Meeting in Ljubljana, Autumn 2018) is to develop a separate / specific schema based on ERML with extensions mainly because ERML is well known by providers and thus easier to use. It should be built as an extension of the Minerals4EU data model as a part of EGDI database.

2 DATA MODELS

A data model (DM) will be used (i) to harvest National Providers aggregated datasets, (ii) to build a Harvesting database and (iii) to provide validated data to BRGM (from BGS).

The ORAMA recommendation is to develop a single data model managing production, resources, reserves and exploration data at the same time. Trade (i.e., import and export) data will not be harvested as it is purchased by BGS, but as it has to be delivered to the Minerals4EU Diffusion Platform.

To build the new e-Minerals Yearbook data model the following list briefly summarizes the main properties to be considered (based on BGS templates used in the Minerals4EU H2020 project):

- Country
- Year
- Commodity (will necessitate a mapping between INSPIRE MR/CGI CommodityCodeValues and BGS list of commodities, with a probable loss of precision, BGS list being more detailed for some commodities)
- Classification System (JORC, NI43-101, UNFC…)
- Classification sub-categories (inferred, indicated, measured, probable, proven, UNFC code…)
- Quantity (value + unit)
- Statistical Type (or something equivalent, which can take the following values)
    - Production
    - Reserve
    - Resource
    - Exploration
        - Number of active licences

        - Number of licences issued
        - Number of companies exploring
        - The area covered by exploration licences
        - The amount of expenditure incurred
    - Trade (data not harvested, only provided by BGS for diffusion)
        - Import data
        - Export data
- Note/Comment

The data model will be built by GEUS.

3 HARVESTING SYSTEM

The classical harvesting procedure can work with national providers such as Geological Surveys (possibly not all). If other institutions are in charge of gathering this aggregated data at national level then, instead of a WFS**, GeoZS will set up a dedicated Web application where each data provider will be able to enter data for their country.**

Harvesting will be done using the INSPIRE code list of commodities. **A mapping is needed at some stage** in order to transfer harvested data to BGS (the e-MYB has its own list of commodities which is different from the INSPIRE one).

GeoZS will establish a harvesting reporting system as part of quality assurance procedures to ensure that data harvesting happens correctly. **Also, a Web application will be set up where data providers will be able to the review their own harvested data by country and commodities.**

4 VERIFICATION AND UPDATES OF RESOURCES, RESERVES AND EXPLORATION DATA IN E-MINERALS YEARBOOK

GeoZS will set up a Web interface, where data on **resources, reserves and exploration** will be visible per country. Through assigned passwords, each data provider will be able to view and enter new data for their country. Only GeoZS and BGS will have administrative rights. The work process will be a sequence of six sub-steps:

- Step 1: Update the list of data providers (BGS will provide existing contact list);
- Step 2: Development of a revised BGS questionnaire:
- Step 3: Setting up the Web interface
- Step 4: Implementation of the revised questionnaire and data entries for 2019 according to the revised questionnaire, conducted through the Web interface.
- Step 5: Quality control of the data entries for 2019, led by BGS because of their extensive involvement in setting up the e-MYB from the beginning.
- Step 6: Update of the European Minerals Yearbook.

5 DELIVERY OF HARVESTED DATA TO BGS

The ORAMA recommendation is to deliver a dump of the harvesting DB to BGS. Such a dump would then contain data related to resources, reserves, exploration and production (if production data is available).

The dump will be a SQL dump. The content of a SQL dump is a large collection of SQL commands in ASCII. Running the script will recreate the PostGreSQL database in the same state as it was when the dump was created.

6 STORAGE, CONTROL AND HARMONISATION OF DATA BY BGS

From the dump, data will follow two different ways:

- Production data will be extracted from the dump and incorporated into the BGS World Mineral Statistics DB and then extracted to feed the BGS website DB.
- Exploration, resource and reserve data would be stored in a BGS' copy of the GeoZS' Harvesting database.

Only the production data will go through BGS' usual quality control and standardisation procedures. Resources and reserves data will not be 'standardised' (nor 'harmonised') because that needs to be done by the data providers. This is because the harmonising process needs additional information that BGS will not have access to. BGS will however do some quality control work and ensure it is presented in an appropriate way.

7 TRANSFER OF VALIDATED/HARMONIZED DATA TO THE DIFFUSION DB AT BRGM

- Production & trade data: BGS already delivers production and trade data via its website (and will continue to do so). This step is about how BGS makes that data available to BRGM for the e-Minerals Yearbook part of the Minerals4EU database which is also the EGDI database. BRGM have indicated that BGS' existing 'Web service' is not suitable. This Web service, which is available at http://www.bgs.ac.uk/mineralsuk/statistics/wms.cfc?method=searchWMS, proposes to visualize the data directly on the web site or to download them in Excel format (xlsx). This Web service can be used to retrieve in Excel format the data for the production, import and export for a maximum of 10 years from 1970 to 2016 (depending on the commodities). The major problem with this service is the Excel output format which needs to be converted to database format (the use of JSON for example would have made the operation easier) before being used to feed the EGDI database, and more problematic, the fact that this service is not designed to retrieve all the data in this way.
  In order to bypass this series of problems, the best solution will be to use a Web service based on the e-Minerals Yearbook data model instead of the BGS website Mineral Statistics service or database dump.

- Resources/reserves/exploration data: the development of a Web service based on the e-Minerals Yearbook data model is needed.

8 REFERENCES

Cassard D., Tertre F., Heijboer T., Schjøth F., Šinigoj J., Sőrés L., Eloranta T., Bide T., Baldé K., Huisman J., Vidal-Legaz B., Mancini L. (in prep.). ORAMA Project, Compatibility of improved datasets with the INSPIRE Directive and existing data models, and identification of necessary evolutions. Deliverable D3.1, Public Document.