MINERAL INTELLIGENCE
FOR EUROPE

Authors and affiliation:
Frands Schjøth and Marianne B. Wiese
[GEUS]

E-mail of lead author:
fsc@geus.dk

Deliverable

Data Management Plan

Version: 30-January-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	D1.1	
Dissemination level	Public	
Deliverable name	Data Management Plan	
Work package	WP1, Management, communication, dissemination and sustainability	
Lead WP/Deliverable beneficiary	GEUS	
Deliverable status		
Submitted (Author(s))	17/12-2018	Frands Schjøth
Verified (WP leader)	30/01/2019	Jørgen Tulstrup
Approved (Coordinator)	30/01/2019	Jørgen Tulstrup

Change	Date	Version
Initial Plan, GEUS	Dec. 2018	0.7
BRGM review	11/01/2019	
BGS supplemented	14/01/2019	
GEUS	30/01/2019	0.8
Data model for MYB		
Decisions taken		

GENERAL INTRODUCTION

This report describes the Data Management Plan (DMP) for the Mintell4EU project. The purpose of the plan is to ensure that research data generated in the projects will be findable, accessible, interoperable and reusable (FAIR).

This DMP is structured according to the Horizon 2020 FAIR Data Management Plan template.

TABLE OF CONTENTS

ACRONYMS

Acronym	Explanation
API	Application Programming Interface
CGI	Commission for the management and Application of Geoscience Information
EGDI	EGDI is EuroGeoSurveys' European Geological Data Infrastructure
EGS	EuroGeoSurveys
e-MYB	electronic Minerals Yearbook
e-MineralsYearbook	electronic Minerals Yearbook
ERML	EarthResourceML is an XML-based data transfer standard for the exchange of digital information for mineral occurrences, mines and mining activity. The model describes the geological features of mineral occurrences, their commodities, mineral resources and reserves. It is also able to describe mines and mining activities, and the production of concentrates, refined products, and waste materials
ERML-Lite	ERML-Lite is an abridged version of the full EarthResourceML model and can be used to deliver simplified views on mineral occurrences and their commodities, mines, mining activities and mine waste products
EURMKB	European Union Raw Materials Knowledge Base
EURare	Development of a sustainable exploitation scheme for Europe's Rare Earth ore deposits
EuroGeoSurveys	The Geological Surveys of Europe, is a not-for-profit organisation representing 37 National Geological Surveys and some regional Surveys in Europe, an overall workforce of several thousand experts
FAIR	Findable, Accessible, Interoperable and Reusable (FAIR) as defined in the Guidelines on FAIR Data Management in Horizon 2020.
GeoERA	Establishing the European Geological Surveys Research Area to deliver a Geological Service for Europe (GeoERA)

Acronym	Explanation
GeoSciML	An XML–based 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
GIP-P	GeoERA Information Platform project → EGDI Information Platform
GeoSciML	GeoSciML is a data model and data transfer standard for geological data - from basic map data to complex relational geological databases.
INSPIRE	Infrastructure for spatial information in Europe, laid down in EU directive 14.03.2007WFS
M4EU	M4EU is used as an acronym used for describing the DB and the data model developed inside the project: the INSPIRE database developed using Unified Modelling Language (UML) the M4EU DataBase (M4EU DB) and providing data to the M4EU DB. The following projects contributed to the development of the M4EU DB: • Minerals4EU, EURARE & ProSUM Mining Waste
Minerals4EU	Minerals Intelligence Network for Europe EU-FP7 project
Mintell4EU	Mineral Intelligence for Europe H2020 project
ProSUM Mining Waste	Prospecting Secondary raw materials in the Urban mine and Mining waste – the mining waste part
RMIS	EU Science Hub on Raw Materials Information Systems developed by JRC
UML	Unified Model Language
UNFC	United Nations Framework Classification for Fossil Energy and Mineral Reserves and Resources 2009 Incorporating Specifications for Its Application
WFS	Web Feature Service, an interface allowing requests for geographical features across the web using platform-independent calls.

1 INTRODUCTION

The overall aim of the Mintell4EU project is to improve the European Union Raw Materials Knowledge Base (EURMKB) by extending the spatial coverage and quality of data from past and ongoing European projects on raw materials (Annex I) and by updating the electronic Minerals Yearbook (e-MineralsYearbook or e-MYB) produced in the Minerals4EU project. The e-MineralsYearbook will be developed in a separate schema in the M4EU database through the Mintell4EU project.

All results from the Mintell4EU project will be integrated in the European Geological Data Infrastructure (EGDI) through the GeoERA Information Platform project (GIP-P), just like geospatial results from all other GeoERA project will. Mintell4EU will extend EGDI, a platform for dissemination of up-to-date, comprehensive, reliable and harmonised raw materials intelligence from Europe. EGDI will, when Mintell4EU is finalised, comprise both a dedicated web portal and an Application Programming Interface (API) to deliver data to the Raw Materials Information System (RMIS) of the European Union.

As part of organising the data management, GeoERA will require that all data is made findable, accessible, interoperable and reusable (FAIR) as defined in the Guidelines on FAIR Data Management in Horizon 2020. This document describes how Mintell4EU plans to manage data in a FAIR way and, being a living document, it will be revised according to the timetable for updates.

2 DATA MANAGEMENT PLANS – MINTELL4EU

2.1 Data summary

2.1.1 What is the purpose of the data collection/generation and its relation to the objectives of the programme?

The goal of Mintell4EU is to improve the European Union Raw Materials Knowledge Base (EURMKB) to facilitate the goals of ensuring security of supply, good environmental management, resource exploitation efficiency and sustainable management for the raw materials sector in Europe. Building upon data collections from previous EU projects, Mintell4EU will extend existing collections with recent data and extend the spatial coverage and improve data quality. Mintell4EU furthermore aims at establishing a fully operational and reliable Database Management System, with routines to keep data comprehensive and up-to-date.

2.1.2 What types and formats of data will the programme generate/collect?

Mintell4EU will particularly extend and improve two distinct data collections. Firstly, the Minerals Inventory comprising mineral occurrences, commodities, mines, reserves and resources as well as mining wastes.

Secondly, the e-MineralsYearbook comprising the following statistical data:

- Mineral production – for a wide range of commodities, for 40 countries, for all years from 2004 to 2019;
- Mineral trade (imports and exports) – for a wide range of commodities, for 36-40 countries, for all years from 2004 to 2018; and
- Mineral resources, reserves and exploration – for a wide range of commodities, up to 40 countries, for the years 2013 and 2019 (but not the intervening period).

The Mineral Inventory data is stored in a relational database, structured according to the EarthResourceML data model. The e-MineralsYearbook will however possibly be in a separate schema but also based on ERML-Lite.

2.1.3 Will you re-use any existing data and how?

The Mineral Inventory is a compilation of data currently stored in databases maintained by national and regional geological survey organisations. Within a previous EU Project, Minerals4EU, the partners agreed on a common INSPIRE-based data model and each partner have mapped their data to this model and exposed it via WFS services. A central relational database is regularly updated with current data from all partners.

The e-MineralsYearbook is a compilation of data from a wide variety of sources which vary according to data type. Some of these data are published by other organisations in different formats but most are not. The production data are published annually by British Geological Survey (BGS) as part of its “World Mineral Production” series of publications. European trade (imports and exports) data up to 2014 were previously published by BGS in a separate series of publications called “European Mineral

Statistics" (now ceased). Resources, reserves and exploration data or information may occasionally be published by national geological surveys but in a variety of forms.

2.1.4 What is the origin of the data?

Data of the Minerals Inventory originates from national and regional geological survey organisations within Europe.

Data in the e-MineralsYearbook originates from a wide variety of sources which vary according to data type. Production data are collected by the British Geological Survey (BGS) on an annual basis from government departments, geological surveys, statistical offices, trade associations, companies and other relevant sources as required. Trade data are purchased in bulk from an agency, assessed for completeness and accuracy by experienced BGS staff and supplemented with data, as required, from the United Nations Commodity Trade database, Eurostat or national statistical agencies. Resources, reserves and exploration data will be collected from national geological surveys using questionnaires.

2.1.5 What is the expected size of the data?

The size of the data is within a few hundred GB.

2.1.6 To whom might it be useful ('data utility')?

Data can be used by public and private decision makers, the raw materials industry and society in general. Data can be used to produce geological maps on mineralized areas and to report exploration and mineral productions, flows and foresight studies.

The improved dataset will be accessible through the EDGI information platform, European Union Raw Materials Knowledge Base (EURMKB) and EU Science Hub on Raw Materials Information Systems (RMIS)

2.2 Fair data

2.2.1 Making data findable, including provisions for metadata

2.2.1.1 Are the data produced and/or used in the programme discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

Metadata for the Minerals Inventory will be made available through the EGDI Portal. The metadata will link to a presentation of the data.

The ORAMA project has decided on the 2nd consortium meeting 12-13 November 2018 to develop guidelines on how to use unique INSPIRE ID's on record level. Mintell4EU will adopt these guidelines.

2.2.1.2 What naming conventions do you follow

In the Minerals Inventory, the M4EU database model is based on the INSPIRE Mineral Resources core and extension xsd schemas, which include elements from other xml schemas. The INSPIRE MRcore together with the extension schema are almost the twin of the ERML mineral resource schemas and contain references to the same code list values. When a data provider creates an INSPIRE WFS or download service, approximately the same database mapping could be used to create an ERML compliant WFS service. Currently, the M4EU setup only creates an INSPIRE WFS service. The following naming conventions are used in the M4EU database:

- Generally, each aspect of the INSPIRE xml schema has a related table or column in the database.
- INSPIRE code lists will be used on record level
    - Currently there is a discrepancy between M4EU and INSPIRE, but this will be taken care of first half of 2019
- On EU level it looks like that additional codes are possibly necessary.
    - Suggestion of new codes to INSPIRE codelist has to follow the INSPIRE guidelines for applying for additional codes to the list
    - (GeoERA continue the discussion with Mintell4EU – FRAME – MINDeSea: How to support and update codelist?)

2.2.1.3 *Outline the approach towards search keywords*

Keywords for metadata will as far as possible be taken from thesauri used by INSPIRE.

2.2.1.4 *Outline the approach for clear versioning*

Partners are encouraged to provide the latest version of the common database model and the latest data. The harvesting database model is always the latest version. The data from each partner for The Minerals Inventory is harvested each month. Data may change with each harvesting. Each record may have the attributes of origin (harvesting url) and date of latest harvesting (pr. country/survey) and database model version.
The ORAMA project will produce a guideline for this and Mintell4EU will follow it.

The production data in e-MineralsYearbook will be recopied from the BGS web database each year in order to capture the inevitable revisions to previous years as well as adding each additional year. The other data types will be added each year with no updates to prior years being expected. The mechanism for this data transfer is being developed under the ORAMA project and Mintell4EU will use this mechanism.

2.2.2 Making data openly accessible

2.2.2.1 *Which data produced and/or used in the programme will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.*

The data in the Minerals Inventory and the e-MineralsYearbook will be made openly available to explore and view. See also next point, regarding download.

2.2.2.2 *Will the data be made accessible through the Information Platform? If not, explain why.*

A download services for data in the Minerals Inventory (the harvested database) will be made available from the EGDI Portal. In the beginning of 2019 it will be examined whether all the data providers will allow this.

The data in the e-MineralsYearbook will be openly available and possibly downloadable.

2.2.2.3 *What methods or software tools are needed to access the data?*

As Mintell4EU Data will be accessible as OGC-services, specific tools will not necessary to access the data.

2.2.2.4 *Is documentation about the software needed to access the data included?*

Irrelevant. See 2.2.2.3.

2.2.2.5 *Is it possible to include the relevant software (e.g. in open source code)?*

Irrelevant. See 2.2.2.3.

2.2.2.6 *Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

Data:
The Minerals Inventory is updated regularly by the Geological Survey of Slovenia (GeoZS), who is harvesting data from each national data provider. Copies of the Minerals Inventory are sent to GEUS and BRGM and will be accessible via the EGDI Portal. As described in this section 2.2.2.2 download of data will also be provided.
The e-MineralsYearbook data will be stored at BGS, GEUS and BRGM.

Metadata:
The national datasets from which the Minerals Inventory is harvested are registered in MICKA metadata catalogue (http://micka.bnhelp.cz/, hosted by the Geological Survey the Czech Republic (CGS)). MICKA is also accessible through the EGDI portal.

The datasets available in the e-MineralsYearbook are (production/trade data) or will be (resources/reserves/exploration data) described in the BGS Discovery Metadata catalogue (https://www.bgs.ac.uk/discoveryMetadata/)

Documentation and codes:

The M4EU UML, DB model, guidelines, WFS service implementation are in the repository http://data.geus.dk/svn/m4eu, which is openly accessible.

2.2.3 Making data interoperable

2.2.3.1 Are the data produced in the programme interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

The Minerals Inventory data is interoperable because the data from regional and national survey organisations are stored in a common database. The interoperability lies in the fact, that each survey is mapping its data into a scheme modelled in accordance with EarthResourcesML. The geospatial data is accessible in a standard format via OGC web services.

The e-MineralsYearbook data will follow the ERML-Lite.

2.2.3.2 What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Metadata will follow INSPIRE guidelines for metadata, which are built upon ISO 19115.

2.2.3.3 Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

The Minerals Inventory uses the official INSPIRE code lists as far as it is possible, see 2.2.1.2

- (GeoERA continue the discussion with Mintell4EU – FRAME – MINDeSea: How to support and update codelist?)

However, in certain cases it will be necessary to extend these vocabularies, to retain fine-grained information. For example to distinguish flaky graphite from solid graphite, and only one term ‘graphite’ is present in the INSPIRE code-list.

- INSPIRE registry http://inspire.ec.europa.eu/codelist
- Extended INSPIRE code list http://data.geus.dk/svn/m4eu/09%20DB%20M4EU/ spreadsheet with all INSPIRE registry and extended where the INSPIRE registry allows extension (at version 1.1.2 at the moment)
- INSPIRE follows ERML and CGI code lists, which also are subject for version maintenance.

2.2.3.4 *In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

See 2.2.3.3.

2.2.4 Increase data re-use (through clarifying licences)

2.2.4.1 *How will the data be licensed to permit the widest re-use possible?*

As described in 2.2.2.2, Mintell4EU plan to provide download facilities for the Minerals Inventory and the partners will have to agree on a clarifying license.

Suggestion for a disclaimer, allowed use & IPR of the portal, see following:

DISCLAIMER

Use of any data, information, or service supplied by EGDI is at the users own risk. EGDI and the rightholders give no warranty, condition or representation as to the quality, accuracy or completeness of the data, information, or service, or its suitability for any use or purpose. All implied conditions relating to the quality or suitability of the data, information, or service and all liabilities arising from the supply of the data, service, or information (including any liability arising in negligence) are excluded to the fullest extent permitted by law.

The EGDI Portal

Allowed use

The use of the EGDI Portal is free of charge for any use, including public, private and commercial use, according to the license.

IPR - General

Any and all Intellectual Property Rights in the geological data provided by the EGDI portal are and shall remain the exclusive property of their respective right holders.

Any form of use / or reproduction must be accompanied by the following acknowledgement:

Used and/or reproduced with the permission of the EGDI project. All Rights Reserved.

For the e-MineralsYearbook, similar considerations will also be described in a future version of this document.

2.2.4.2 *When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

The data providers at each partner side decides when to publish data. They publish to their own national/regional database, and map data into a Minerals4EU common database, which is in turn harvested to the Minerals Inventory. For the e-MYB, data are

confidential until they have been checked for quality control purposes, which happens before they are uploaded to the e-MYB.

2.2.4.3 Are the data produced and/or used in the project useable by third parties, in particular after the end of the programme? If the re-use of some data is restricted, explain why.

Data in the Minerals Inventory will also be available after the end of the project, since the Mintell4EU project plans to implement a Database Management System, and ensure the continued operation, maintenance and accessibility of the database.

2.2.4.4 How long is it intended that the data remains re-usable?

See 2.2.4.3

2.2.4.5 Are data quality assurance processes described?

The quality of the data depends on each partners quality assurance procedures. The Mintell4EU project is not aware of the procedures at each Geological Survey, but recommendations for QC procedures will be sent to the data providers in coordination with the ORAMA project.

For the e-MYB the data are subjected to stringent quality control procedures used by BGS for many decades. These procedures are not currently published.

2.3 Allocation of resources

2.3.1 What are the costs for making data FAIR in your project?

Data from the Minerals Inventory and the e-MineralsYearbook were made findable by previous EU projects. Making these data accessible / downloadable, interoperable and reusable will require some work:

a. Partners must meet and agree whether all harvested data can be made downloadable. If any part of the data cannot be made accessible, it must be explained why. The cost for this work is low.
b. The EGDI platform must set up new services to download data. The cost for this work is low.
c. Data managers must suggest necessary code-list changes to official code-list registries, first and foremost INSPIRE registry and CGI registry. The cost of negotiating code list changes with the relevant authorities is medium.
d. Data managers must suggest proper license / copyright for download services. The cost of this work is low.

2.3.2 Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

See 2.2.4.3

2.4 Data security

2.4.1 What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

The Minerals Inventory stored in the central EGDI database is currently secured through the fact that it is operated by BRGM and included in their operational procedures. A copy of the database is currently at GEUS. The database for the e-MineralsYearbook will be included in the same database(s).

2.4.2 Is the data safely stored in certified repositories for long term preservation and curation?

Currently GeoZS, BRGM and GEUS have the latest backup of the harvested data. The harvesting is on monthly base.

2.5 Other issues

2.5.1 Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

So far: No.

ANNEX I

Background info Minerals4EU, ProSUM urban mine & RESEERVE

M4EU

M4EU is used as an acronym used for describing the DB and the data model developed inside the project: the INSPIRE database developed using Unified Modelling Language (UML) the M4EU DataBase (M4EU DB) and providing data to the M4EU DB.

The following projects contributed to the development of the M4EU DB:

- Minerals4EU, EURARE & ProSUM Mining Waste

The partners involved in Minerals4EU is on country level and are mostly geological surveys. Each partner contributes with individual spatial information on raw material and secondary raw material including mining waste. Each partner is responsible on updating the M4EU DB through their own local services and each partners local service is harvested (done by Geo-ZS) each month and sent to Minerals4EU WEB-portal at BRGM

The Minerals4EU Mineral Yearbook is from 2013 without facility in updating the yearbook and aggregated to country level in Europe

At the moment there is 23 partners in Minerals4EU.

- Minerals4EU 23 partners is
    - Austria, Belgium, Croatia, Cyprus, Czech Republic, Denmark, Finland, France, Greece, Hungary, Ireland, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Ukraine & United Kingdom
- None Minerals4EU partners
    - Albania, Andorra, Belarus, Bosnia and Herzegovina, Bulgaria, Estonia, Faroe Islands, Germany, Gibraltar, Iceland, Isle of Man, Italy, Kosovo, Latvia, Lichtenstein, Lithuania, Luxembourg, Macedonia, Malta, Moldova, Monaco, Montenegro, Russia, San Marino, Serbia, Turkey & Vatican city

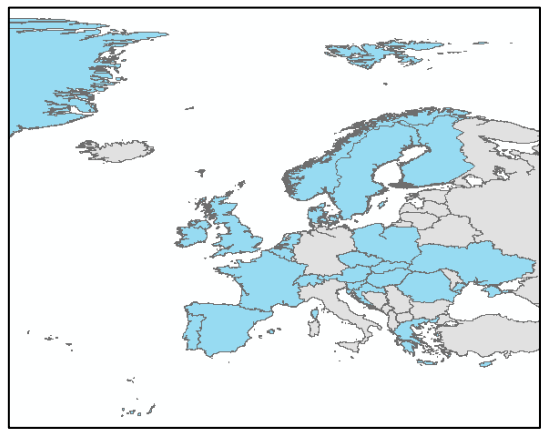

Figure 1: M4EU 23 partners in blue color

ProSUM Urban Mine

The ProSUM Urban Mine part worked on Batteries, End-of-Life Vehicle (ELV) & Waste Electrical and Electronic Equipment (WEEE) on country level in Europe.

RESEERVE partners

- Albania, Bosnia and Herzegovina, Croatia, Macedonia, Montenegro and Serbia

Figure 2: RESEERVE partners in green color

The project's hope is to get at least 2 partners into Minerals4EU partnership

ANNEX II

2.3.1.1 Current status: How will the data be licensed to permit the widest re-use possible?

Each of the 23 M4EU partners provides WFS

- Minerals4EU
- EURare
- ProSUM mining waste

In which by default data can be downloaded as XML-files. The portal data is viewable and query able but no download is provided. The metadata can be downloaded.

ProSUM Urban Mine data is viewable, queryable and in 2019 a download facility will be implemented in the web-browser.