AfricaMaVal

Coordination and Support Action (CSA)

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or he European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

Start date : 2022-06-01 Duration : 42 Months

INSPIRE compliant data set of Pan-African inventory of the known CRM deposits

Authors : Dr. Daniel OLIVEIRA (LNEG), Carla Fortes, Daniel de Oliveira, Cátia Prazeres

AfricaMaVal - Contract Number: 101057832

Project officer: Victoria Leroy

Document title	INSPIRE compliant data set of Pan-African inventory of the known CRM deposits
Author(s)	Dr. Daniel OLIVEIRA, Carla Fortes, Daniel de Oliveira, Cátia Prazeres
Number of pages	31
Document type	Deliverable
Work Package	WP1
Document number	D1.1
Issued by	LNEG
Date of completion	2023-04-13 16:12:39
Dissemination level	Public

Summary

Providing an INSPIRE-compliant accurate (as far as possible) inventory of: (i) the non-energy, mineral raw materials deposits and where possible distinguishing between active and abandoned deposits; and (ii) incorporating, where possible, available mineral resource/reserve data.

Approval

Date	By
2023-04-14 09:26:42	Mrs. Carol ZAMMIT (BRGM)

Horizon Europe Framework Programme (HORIZON)

D1.1 – INSPIRE compliant data set of Pan-African inventory of the known CRM deposits

WP1 - Task 1.1

February 2023

Carla Fortes[1], Daniel de Oliveira[1], Cátia Prazeres[1]

[1]Laboratório Nacional de Energia e Geologia (I.P.)

Funded by
the European Union

Disclaimer

The content of this deliverable reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Funded by
the European Union

Document information

Grant Agreement / Proposal ID	101057832
Project Title	EU-**Africa** Partnership on Raw **Ma**terial **Val**ue chains
Project Acronym	AfricaMaVal
Project Coordinator	Guillaneau Jean-Claude (jc.guillaneau@brgm.fr) - BRGM
Project starting date (duration)	1st June 2022 (42 months)
Related Work Package	WP1
Related Task(s)	Task 1.1
Lead Organisation	Laboratório Nacional de Energia e Geologia (I.P.)
Contributing Partner(s)	BRGM
Due Date	January 2023
Submission Date	February 2023
Dissemination level	

History

Date	Version	Submitted by	Reviewed by	Comments
23/01/2023	Draft	LNEG	BRGM	
17/02/2023	Draft	LNEG	BRGM	

Table of Contents

1 Introductory statement for WP1 9
1.1 Objectives of Task 1.1 of WP1 9
1.2 The concept of ECRM 10
2 INSPIRE-COMPLIANT MINERALS DATABASES 10
2.1 The state-of-the-art in AfricaMaVal 11
2.1.1 AfricaMaVal compromise/methodology 17
2.1.2 Data harmonization 20
2.1.3 Data confidentiality 22
2.1.4 Data management and visualization 22
3 Overview of the final database 23
1. Annex 1 - Minerals4EU-WP5: The conceptual model for M4EU Database 30

List of Figures

Figure 1 – European minerals databases and intelligence networks. 11
Figure 2 – Graphical overview of the whole Minerals4EU data model. This complex model was developed by TNO and GEUS and was based on the INSPIRE MR V. 3.0 data model, on the ERML v. 2.0 data model and integrating GeoSciML "extensions" (Waardenburg et al., 2014). Please see Annex 1 for legible copy. 12
Figure 3 - Overview of the MineralOccurrence, OreMeasure and commodity relation to the Minerals4EU data model (Waardenburg et al., 2014) 14
Figure 4 - Overview of the Mine, MiningActivity, OreMeasure relation to the Minerals4EU data model (Waardenburg et al., 2014). 15

List of Tables

Table 1 - Tables that make up the mandatory INSPIRE nucleus of mineral resources. 13
Table 2 - INSPIRE – Definition of MineStatusType: Values indicating the operational status of the mine. 16

Funded by
the European Union

Table 3 - The complete summary of the vocabulary used (INSPIRE Code lists) in line with the Minerals4EU Data model for AfricaMaVal database version 1.0 19
Table 4 - The complete attributes reference table for AfricaMaVal database version 1.0 24

Abbreviations and Acronyms

Acronym	Description
CRM	Critical Raw Materials
CSRM	Strategic and Critical Raw Materials
EURMKB	European Union Raw Materials Knowledge Base
RMIS	Raw Materials Information System
WP	Work Package
FRAME	Forecasting and Assessing Europe's Strategic Raw Materials needs. www.frame.lneg.pt
EU-MKDP	European Mineral knowledge Data platform
SDI	Spatial Data Infrastructure
USGS	United States Geological Survey
WMS	Web Map Service
WFS	Web Feature Service
CBA	Cell based association method
GDS	Geographic Data Services

Funded by
the European Union

Executive Summary

Work Package (WP) 1 is designed to show the current African CSRM, the available sites for mining and mineral processing and placing this information in the hands of the European political and industrial decision makers to focus on diversifying EU supply chains. The WP deals on an Africa-wide scale while also addressing the regional and local levels via its identification of possible and existing mineral and mining clusters (this task; Task 1.1), ore processing and refining capacities (Task 1.3), as well as artisanal and small-scale mining and processing developments (Task 1.4). Based on task 1.1 it will undertake mineral favourability/predictability mapping (Task 1.2).

Deliverable 1.1, object of this report, deals with the delivery of an INSPIRE-compliant AfricaMaVal geodatabase with the existing African mineral and mining clusters.

Creating a brand-new database from scratch considering the objectives of the project and given the time frame available to do so, was completely unfeasible. Hence, the existing data from LNEG and BRGM was pooled to create the volume necessary to allow for making reliable predictability maps in task 1.2 that follows this one later. Hence, data points from LNEG and BRGM (SigAfrique DB) were merged into one hybrid dataset because some points were INSPIRE-compliant and most not. Therefore, the work carried out for this task involved a considerable volume of data harmonization procedures, which is the process of bringing data from several different sources into alignment so that it can be integrated and used effectively. The processes involved were: A) To identify the sources and define goals (this includes determining the format, structure, and content of the data, identifying the specific data elements that need to be harmonized and the desired format for the harmonized data); B - To optimize time and resources, all information is merged to create a single INSPIRE-compliant dataset that integrates the existing data and will be updated with newer data; C- Because the geodatabase is not static, new data were and are being gathered from independent reports and commercial industry intelligence data sources (e.g., S&P Capital IQ). This integration includes creating a mapping of the data elements from the different sources and determining any necessary data transformations or cleaning that need to be performed. The data cleaning performed involves identifying and removing errors, inconsistencies, duplicate data, correcting spelling mistakes, standardizing units of measurements and removing outliers; and finally, D- result validation.

With the work complete, we can estimate, at this point, that the AfricaMaVal GIS geodatabase will comprise in this early stage of development more than 20000 mineral occurrences and by the time the final geodatabase is handed over, with the addition of further data, it will represent more than 22000 data points.

To overcome confidentiality issues, the final geodatabase will be released as a WMS.

Keywords

AfricaMaVal geodatabase, hybrid data set, INSPIRE-compliant, identify, harmonisation, result validation, web map service.

1 Introductory statement for WP1

This work package is designed to unravel the current knowledge about African CSRM, the available sites for mining and mineral processing and placing this information in the hands of the European political and industrial decision makers. This will contribute to several anticipated project outcomes, with a focus on diversifying EU supply chains while also outlining possible sustainability. The WP deals on an Africa-wide scale while also addressing the regional and local levels via its identification of possible and existing mineral and mining clusters, ore processing and refining capacities, as well as artisanal and small-scale mining and processing developments.

WP1 is divided into four tasks, namely:

- Task 1.1 - provides the base for the subsequent work in this WP, constructing a database of existing known CSRM deposits in Africa.
- Task 1.2 - acknowledges the heterogeneity of some of the datasets and will perform a predictive assessment to reveal hidden trends.
- Task 1.3 - will perform an analysis of the existing ore processing and refining capacities to understand the outlet points of pre-concentrates of the CSRM.
- Task 1.4 - integrates and combines the findings of all previous tasks into one final report.

Dialogue with African and EU partners is variably integrated into individual tasks and supplements dialogue and outreach activities performed in other work packages (WPs 3, 5, 7) as well as through individual country case studies.

1.1 Objectives of Task 1.1 of WP1

The purpose of this work package is to integrate information on the pan-African Critical and Strategic Raw Materials (CSRM) supply potential into one coherent and centralized dataset (e.g., RMIS, EURMKB). The WP focuses on the Pan-African inventory and predictive mapping assessment of the CSRM, and Pan-African inventory of existing ore processing and refining capacities.

The aim of this task is to make a review of published studies and reports that will go towards providing an INSPIRE-compliant accurate inventory of the non-energy, mineral raw materials deposits and where possible distinguishing between active and abandoned deposits by incorporating, where possible, available mineral resource/reserve data. This task will be conducted through a desk study based on available international mining databases (S&P, SIG Afrique) without infringing on copyrighted private or confidential sources, published scientific papers and mining press-releases (e.g., ASX and Sedar-related), and available project

development studies. This inventory allows economical assessment of potential resources related to CRM.

1.2 The concept of ECRM

Minerals and metals defined as strategic to be addressed by the AfricaMaVal project comprise the ones included in the European Critical Raw Materials (CRM) list, the minerals that are used in the Li-battery manufacturing and electric mobility society in general, the minerals needed by the decarbonization targets of the Energy Intensive Industries (EII), the minerals and metals which energy transition and low-carbon technologies are dependent on, and the metals required by the electronics and high-tech industry (Frame, 2018). This extended critical raw materials (ECRM) to be included and targeted in the AfricaMaVal project are: Antimony (Sb), Baryte, Bauxite, Beryllium (Be), Bismuth (Bi), Borate, Cobalt (Co), Copper (Cu), Fluorspar, Gallium (Ga), Germanium (Ge), Hafnium (Hf), Indium (In), Lithium (Li), Magnesium (Mg), Manganese (Mn), Natural graphite, Nickel (Ni), Niobium (Nb), Coking coal, Phosphate rock, Phosphorus (P), Silicon metal (Si), Scandium (Sc), Strontium, Tantalum (Ta), Tin (Sn), Titanium (Ti), Tungsten (W), Vanadium (V),

HREEs-Heavy Rare Earth Elements: (Dysprosium-Dy, Erbium-Er, Europium-Eu, Gadolinium-Gd, Holmium-Ho, Lutetium-Lu, Terbium-Tb, Thulium-Tm, Ytterbium-Yb, Yttrium-Y),

LREEs-Light Rare Earth Elements: (Cerium-Ce, Lanthanum-La, Neodymium-Nd, Praseodymium-Pr, Promethium, Samarium-Sm),

PGM-Platinum Group Metals: (Iridium-Ir, Palladium-Pd, Platinum-Pt, Rhodium-Rh, Ruthenium-Ru).

2 INSPIRE-COMPLIANT MINERALS DATABASES

The issue of compiling INSPIRE - compliant minerals data and making them available to everyone online started almost a decade ago with projects ProMine and EuroGeoSource (Fortes et al., 2013). Evolving from that was the implementation of project Minerals4EU, where a significant leap was made to augment existing data sets and make these more complete and describe more closely the reality in mineral deposits and their variations (e.g., Cassard et al., 2014a, 2014b; Vuollo et al., 2014, Cassard et al., 2015).

The European project Minerals4EU created the European Mineral knowledge Data platform (EU-MKDP), to provide harmonized data related to mineral resources, as well as statistical related information, e.g., the Minerals Yearbook. The proposed technical solutions assured an effective and sustainable system designed for facilitating data updates and maintenance, and for offering

a full and seamless access to information related to the whole mineral resources value chain (Cassard, 2015).

The EU-MKDP represents the first pan European raw materials knowledge base, constituting one of the first bricks of the European geological data infrastructure, in a time where the member states were taking the first steps towards the implementation of the INSPIRE Directive (Cassard et al., 2014a).

The previous work carried out in this domain has laid a solid foundation for the work that is proposed in AfricaMaVal. Several past and ongoing EU projects addressed objectives and issues related to CRM and other strategic minerals at the various levels of details (Fig. 1).

Figure 1 – European minerals databases and intelligence networks.

2.1 The state-of-the-art in AfricaMaVal

In accordance with the definition of the INSPIRE directive (EU/2/2007), the topic of mineral resources (category 21, Annex III) includes in its definition metallic ores, industrial minerals, etc., and when relevant, information on quotas of feature depth/height. The mineral resource data model is organized around two broad categories of information (INSPIRE, 2013):

- the description and location of mines and mining activities and

- the description and location of mineral resources, including their classification, estimates of quantity, as well as a description of the main "commodities" on the market.

The core data model for Mineral Resources, which is based on GeoSciML and EarthResourceML developed by the international geosciences community, in particular Geological Survey Organizations (www.geosciml.org/), provides the main object types and properties requested by all examples of use: the location of mineral resources (Mines and earth resources), the main commodities, and the exploitation type (INSPIRE, 2013).

The Mineral Resources core data model adopted by the AfricaMaVal team was the same as that used by the Minerals4EU team (Fig. 2) and integrates all the tables designated in Table 1 (INSPIRE, 2013).

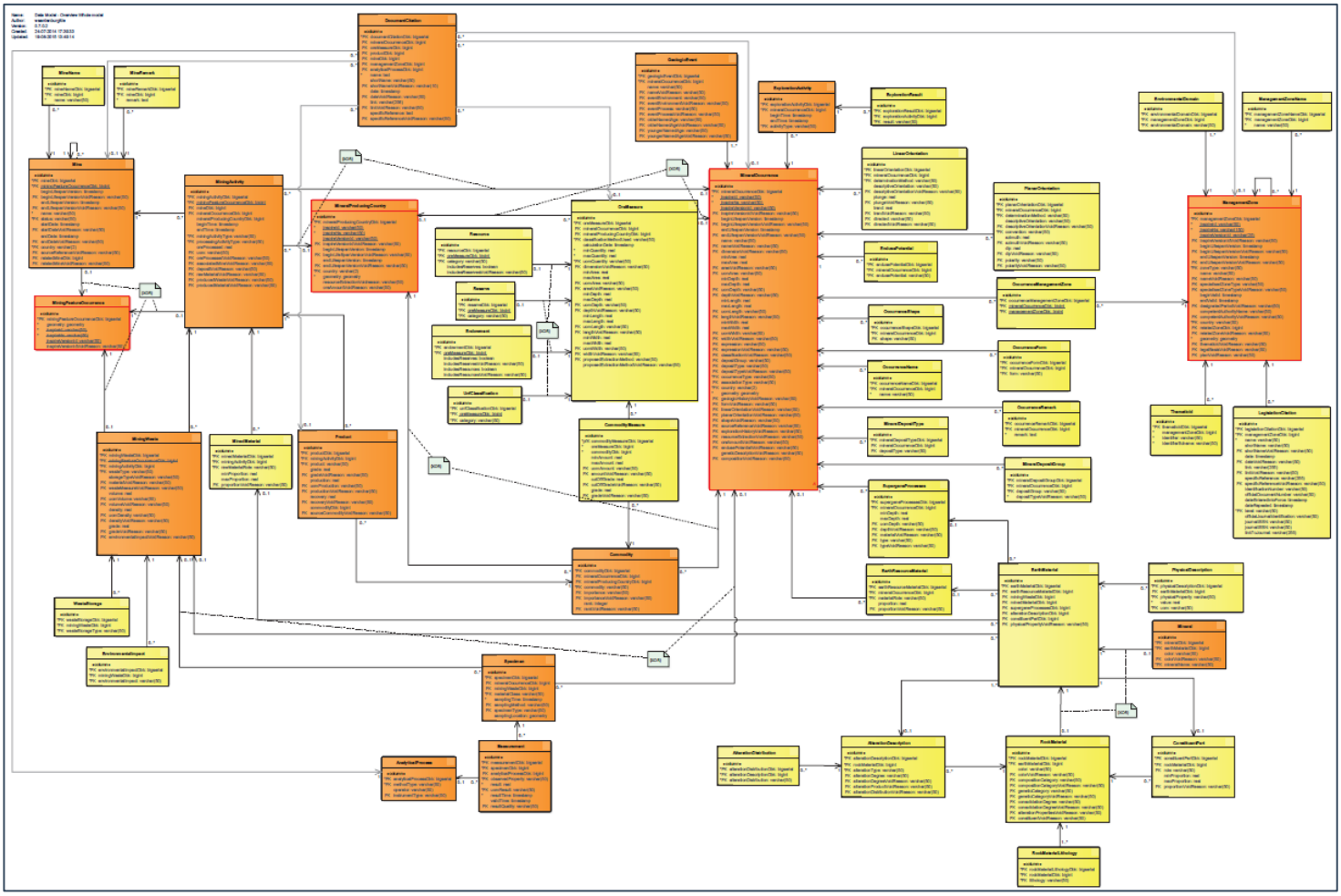


Figure 2 – Graphical overview of the whole Minerals4EU data model. This complex model was developed by TNO and GEUS and was based on the INSPIRE MR V. 3.0 data model, on the ERML v. 2.0 data model and integrating GeoSciML "extensions" (Waardenburg et al., 2014). Please see Annex 1 for legible copy.

Table 1 - Information that integrates the mandatory INSPIRE core of mineral resources.

Mandatory Tables	Attributes associated with the tables
MineralOccurrence	MineralOccurrence
	OccurrenceForm
	OccurrenceShape
	LinearOrientation
	PlanarOrientation
	DocumentCitation
	ExplorationActivity
	Explorationresult
	Geologicevent
MiningActivity	MiningActivity
	MiningFeatureOccurrence
	Mine
	MineName
OreMeasure	OreMeasure
	Reserve
	Resource
	Endowment
	CommodityMeasure
	Commodity

Some of the definitions of the attributes in the table 1 (INSPIRE, 2013) and the Minerals4EU data model (adopted by AfricaMaVal team) (Figs. 3, 4):

- **Mineral Occurrence** - could be a prospect, an occurrence, a mineral deposit, an ore deposit, etc. (but not a lode, a field, a district, or a province).
- **Mine** is an excavation for the extraction of mineral deposits. “True” mines are underground workings and open-pit workings (also called open-sky mines) generally for the extraction of metallic commodities. The Mine spatial object type also includes open workings generally for the extraction of industrial minerals, commonly referred to as quarries.
- **Mining Activity**, related to a Mine, describes the process of extracting metallic or non-metallic mineral deposits from the Earth.
- **Mining Feature Occurrence** is an occurrence of a Mining Feature, it carries some properties and the geometry and/or location.
- The **Ore Measure** is an estimated or calculated amount of ore and grade that exist within an Earth Resource, in terms of its resource, reserve and endowment
- An Earth Resource has an associated **Exploration Activity** to describe the process leading to the discovery and assessment of the resource.
- The **Commodity** describes the material of economic interest in the Earth Resource.
- The **Mining Feature** class represents a conceptual feature that exists coherently in the world and corresponds with a "**Mine**" or a "**Mining Activity**", locatable and identifiable features in time and/or space. The **Mining Feature Occurrence** is an occurrence of a Mining Feature, it carries some properties and the geometry and/or location.

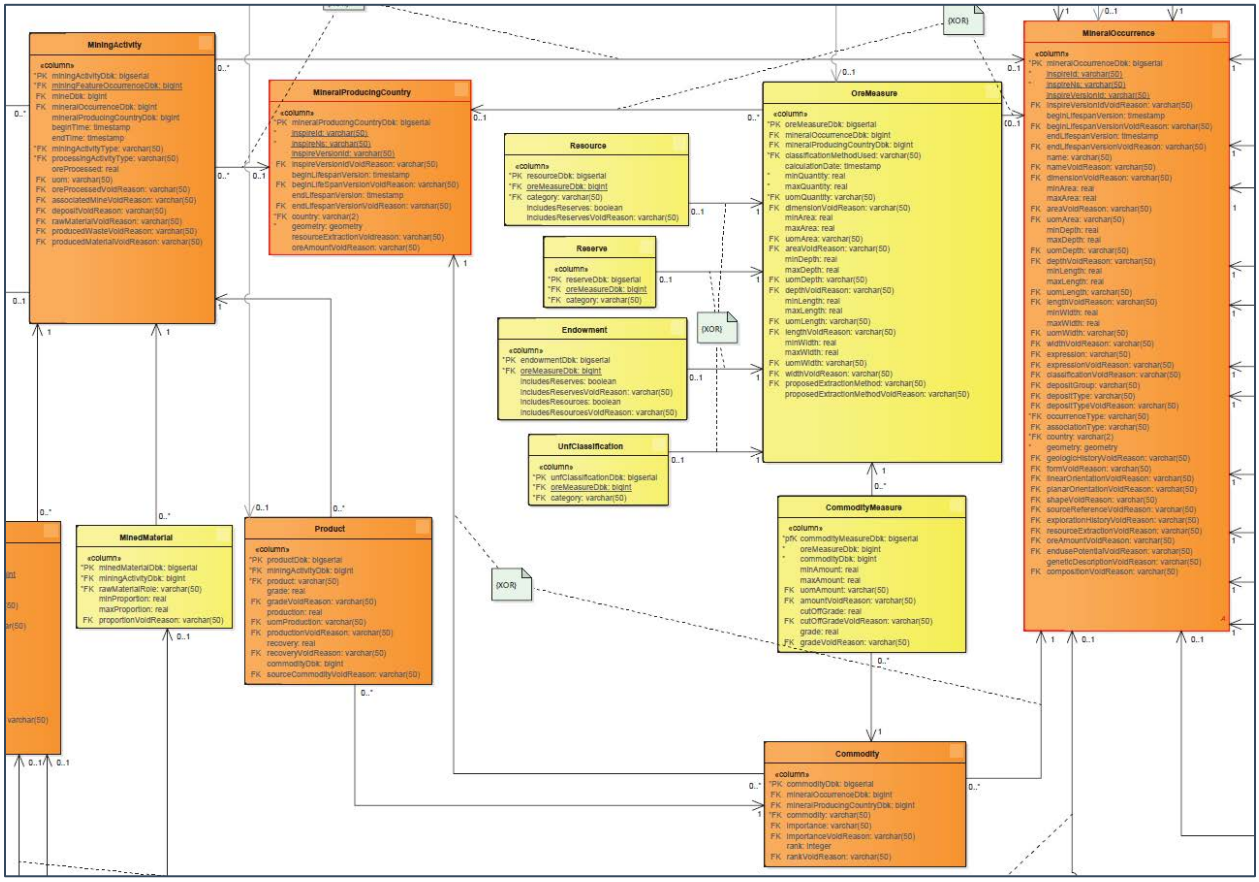


Figure 3 - Overview of the MineralOccurrence, OreMeasure and commodity relation to the Minerals4EU data model (Waardenburg et al., 2014)

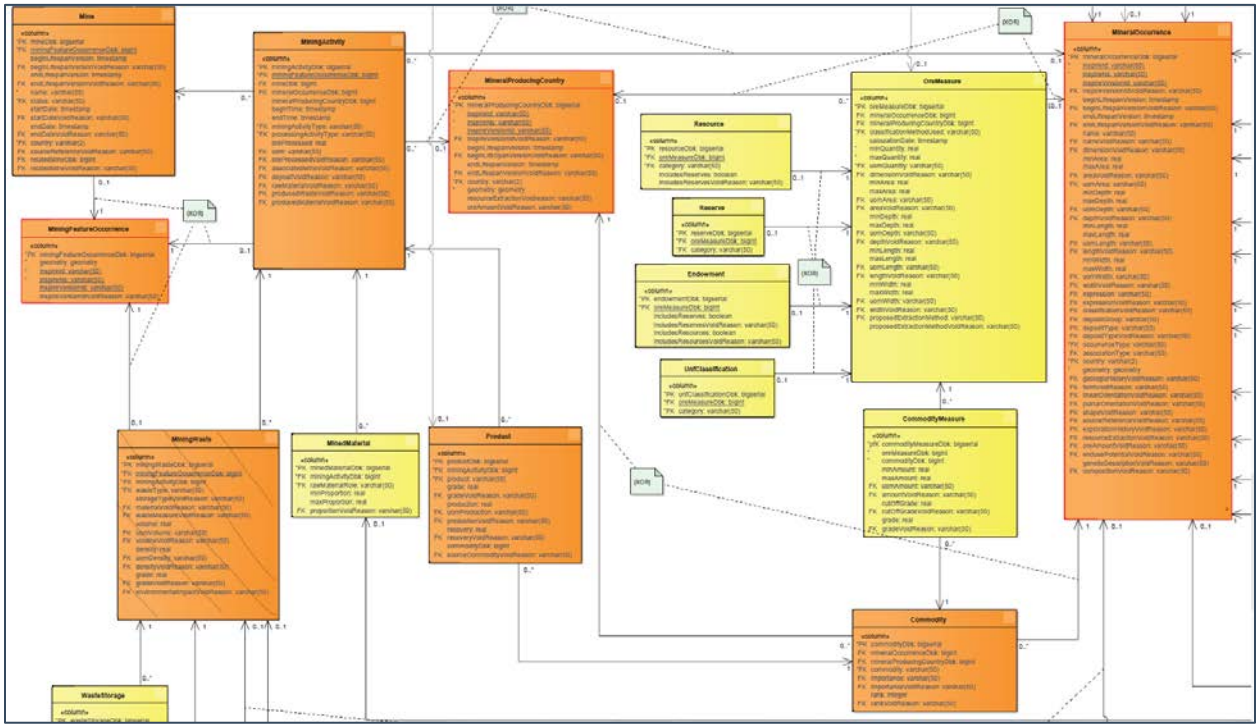


Figure 4 - Overview of the Mine, MiningActivity, OreMeasure relation to the Minerals4EU data model (Waardenburg et al., 2014).

INSPIRE uses the term occurrence type (**OccurrenceTypeType**), from which we have adopted the following vocabulary (INSPIRE, 2013):

Occurrence: An occurrence is a locality with an enrichment above the normal (background) value. There is no measure of economic value or probability of mining with profit. Therefore, our definition may vary from what some other writers use, where an occurrence is a locality with little probability of economic extraction (Goodenough et al., 2016). In our work, every locality is identified as a place where geological processes have resulted in a mineralogical enrichment.

Deposit: A deposit is a locality with an enrichment, which in addition, has an estimation of grade and tonnage. This estimation can be compliant with international standards or noncompliant. We do not infer any economic significance since this will vary with many time variable factors.

Prospect: A prospect is defined according to INSPIRE as an area that is a potential site of mineral deposits, based on preliminary exploration or previous exploration. A prospect usually includes one or more occurrences. Grade and/or tonnage information may or may not be available.

Commodity type: For commodity types, we have used the INSPIRE-defined types; Li, Co, and graphite. Additional precision should be unnecessary.

INSPIRE uses the term mine status (**MineStatusType)**, for which we have used the following vocabulary (INSPIRE, 2013) in table 2:

Table 2 - INSPIRE – Definition of MineStatusType: Values indicating the operational status of the mine.

INSPIRE_codelists			
MineStatus	http://INSPIRE.ec.europa.eu/codelist/MineStatusValue/		
Code	**name**	**Parent**	**description**
operating	operating		A mine is operating.
operatingContinuously	operating continuously	operating	A mine is operating continuously.
operatingIntermittently	operating intermittently	operating	A mine is operating intermittently.
notOperating	not operating		A mine is not operating.
closed	closed	notOperating	A mine can be closed for technical, economical, or technical-economic reasons.
abandoned	abandoned	notOperating	A mine is abandoned.
careAndMaintenance	care and maintenance	notOperating	A mine is under care and maintenance.
retention	retention	notOperating	A mine can be kept unexploited until the price of contained commodity(ies) makes it economical.
historic	historic	notOperating	An 'old' mine which has been exploited before 1900.
underDevelopment	under development		Under development.
construction	under construction	underDevelopment	Under construction.
pendingApproval	pending approval	underDevelopment	A mine waiting for the exploitation authorization, generally given by a State Mining Engineering Department.
feasibility	feasibility	underDevelopment	Technical economic study aimed at assessing the possibility to launching a mine venture.

Funded by
the European Union

2.1.1 AfricaMaVal compromise/methodology

As in Europe, there is awareness in Africa of the need for geospatial information. However, with the European Union being far more advanced in the development of their institutional structures, decision-making processes, policies, and strategies for unification compared to the African Union, it is no wonder that there is a great awareness and understanding of the need for geospatial information in Europe. A further complicating factor is that African countries are far behind most European countries in the development of their geospatial information. This is why SDI is not gaining the recognition that it has in Europe. Nevertheless, there is a glimmer of hope with some new projects and partnerships recognizing the need to establish regional cooperation in the development of geospatial information (Schwabe et al., 2009).

Creating a brand-new database from scratch considering the objectives of the project and given the time available to do so, was completely unfeasible. Hence, the existing data from LNEG and BRGM was pooled to create the volume necessary to allow for making reliable predictability maps in task 1.2 that follows this one later.

However, one important aspect is that the BRGM data (termed SigAfrique) is not INSPIRE-compliant. Therefore, the first step of the work was to create a hybrid data set with INSPIRE-compliant terms and some not INSPIRE-compliant (Fig. 5).

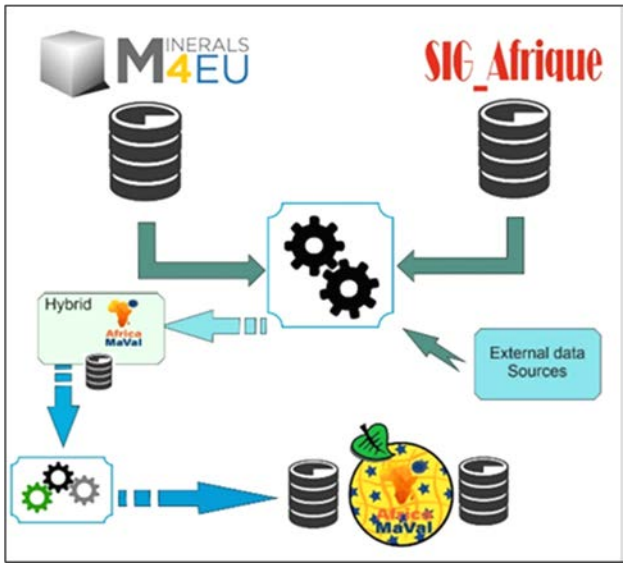


Figure 5 – AfricaMaVal data harmonization for arriving at a AfricaMaVal INSPIRE-compliant DB using a variety of available resources.

Therefore, it would make perfect sense for the AfricaMaVal team, to adopt the INSPIRE directive and consequent projects (Fig. 1) that have outlined a conceptual data model for mineral resources over more than a decade, to establish the Africa's needs: Data on African mineral resources often come from multiple sources, each with its own nomenclature and structure. Also, INPSIRE-compliant code lists are far more simplified than the SigAfrique adopted code lists. Hence, a degree of simplification is needed. For example, for Deposit Type, SIG Afrique will classify a deposit as: "carbonatite-hosted ore deposits: REE, Nb, Ti, Zr, U, Th (P, F, Cu, Ba, Sr, etc." while the INSPIRE code lists will simply use "carbonatite". In total INSPIRE uses a maximum of 49 deposit types while SigAfrique uses 313 different deposit types.

Another example is the Mineral Occurrence type and Mine status type (Fig. 6): In SigAfrique (table A), Mineral Occurrence type and Mine status type are merged into a single table (table A, 72 fields), but, in AfricaMaVal to fulfill all requirements of the INSPIRE directive, we need to split that information in two separate tables. All the data needs to be reclassified and split according to the code list OccurrenceType (Table B, 8 fields) and code list MineStatus (Table C, 14 fields). Therefore, we must establish the relationship between the fields from Tables A, B and C to achieve an automatic insertion into Tables B and C from Table A.

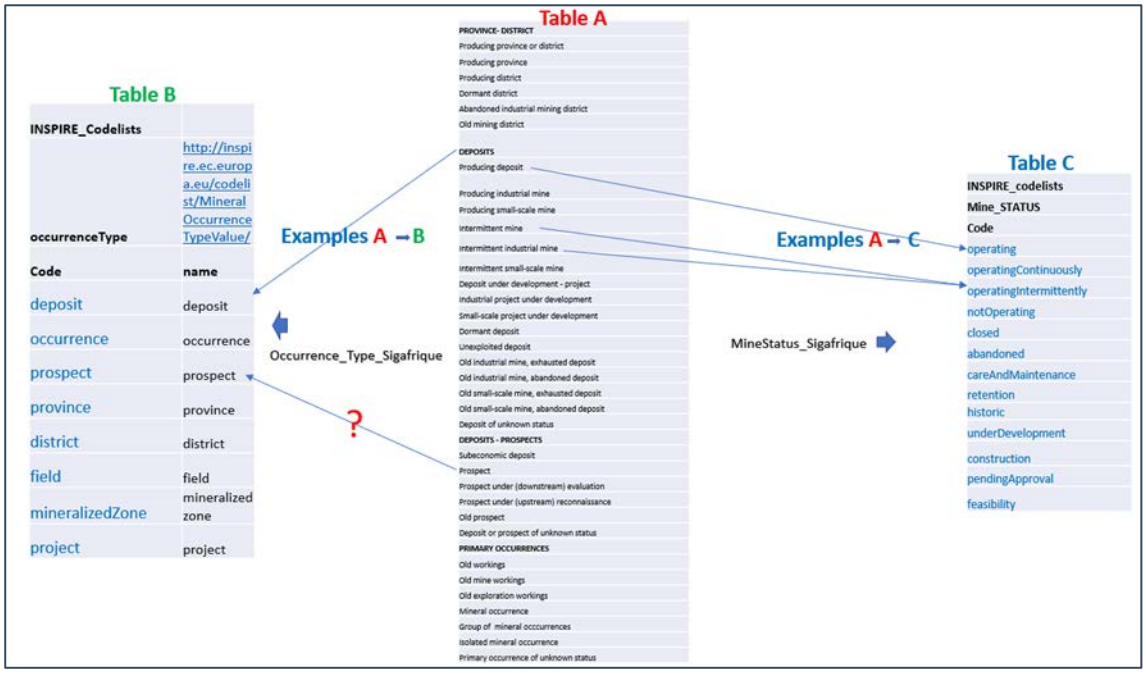


Figure 6- Schematic for the data split from Table A into Tables B (Occurrencetype) and C (MineStatus).

In table 3, we can see all the INSPIRE code lists and AfricaMaVal code lists adopted. This table gives the AfricaMaVal data providers a complete summary of the vocabulary to be used in line with the Minerals4EU Data model, which is attached in the end of this report (Annex I - Code lists).

Table 3 - The complete summary of the vocabulary used (INSPIRE Code lists) in line with the Minerals4EU Data model for AfricaMaVal database version 1.0

INSPIRE Code list name	INSPIRE_URI	AfricaMaVal Code list name adopted in version 1.0
ClassificationMethodUsedValue	http://INSPIRE.ec.europa.eu/codelist/ClassificationMethodUsedValue	INSPIRE Code list adopted
CommodityCodeValue	http://INSPIRE.ec.europa.eu/codelist/CommodityCodeValue	INSPIRE Code list adopted
ExplorationActivityTypeValue	http://INSPIRE.ec.europa.eu/codelist/ExplorationActivityTypeValue	INSPIRE Code list adopted
ImportanceValue	http://INSPIRE.ec.europa.eu/codelist/ImportanceValue	INSPIRE Code list adopted
MineStatusValue	http://INSPIRE.ec.europa.eu/codelist/MineStatusValue	INSPIRE Code list adopted
MineralDepositGroupValue	http://INSPIRE.ec.europa.eu/codelist/MineralDepositGroupValue	INSPIRE Code list adopted
MineralDepositTypeValue	http://INSPIRE.ec.europa.eu/codelist/MineralDepositTypeValue	MineralDepositType_ SigAfrique_Type
MineralOccurrenceTypeValue	http://INSPIRE.ec.europa.eu/codelist/MineralOccurrenceTypeValue	MineralOccurrence_ SigAfrique_Type
MiningActivityTypeValue	http://INSPIRE.ec.europa.eu/codelist/MiningActivityTypeValue	MiningActivityType_ SigAfrique_Type
ProcessingActivityTypeValue	http://INSPIRE.ec.europa.eu/codelist/ProcessingActivityTypeValue	INSPIRE Code list adopted
ReserveCategoryValue	http://INSPIRE.ec.europa.eu/codelist/ReserveCategoryValue	INSPIRE Code list adopted
ResourceCategoryValue	http://INSPIRE.ec.europa.eu/codelist/ResourceCategoryValue	INSPIRE Code list adopted
LithologyValue	http://INSPIRE.ec.europa.eu/codelist/LithologyValue	LithologyValue_ SigAfrique_Type
Earth-resource-material-role	http://resource.geosciml.org/classifier/cgi/earth-resource-material-role	INSPIRE Code list adopted
Earth Resource FormValue	http://resource.geosciml.org/classifier/cgi/earth-resource-form/	INSPIRE Code list adopted
Earth Resource ShapeValue	http://resource.geosciml.org/classifier/cgi/earth-resource-shape/	EarthResourceShape_ SigAfrique_Type
AlterationTypeTerm	http://resource.geosciml.org/classifierscheme/cgi/201012/alterationtype	INSPIRE Code list adopted
MineralNameTerm	http://pubsites.uws.edu.au/ima-cnmnc/IMA_Master_List_(2014-07).pdf	INSPIRE Code list adopted
CountryCode	http://INSPIRE.ec.europa.eu/codelist/CountryCode	INSPIRE Code list adopted

Funded by
the European Union

OreMeasureCategoryType (UNFC)	CGI	INSPIRE Code list adopted
UomVolumeType	M4EU-addition	INSPIRE Code list adopted
UomWeightType	M4EU-addition	INSPIRE Code list adopted

2.1.2 Data harmonization

How to get data from different sources into a consistent, standardized, accurate, and comprehensive format so the data can be consumed and analysed? As data size increases drastically, its variety also increases. Investigating such heterogeneous data is one of the most challenging tasks in information management and data analytics. The heterogeneity and decentralization of data sources affect data visualization and prediction, thereby influencing analytical results accordingly. Data harmonization corresponds to a field that unifies the representation of such a disparate nature of data (Kumar et al., 2021).

Data harmonization is the process of bringing data from several different sources into alignment so that it can be integrated and used effectively. Some processes/methodologies (see below) followed by the AfricaMaVal team for harmonize the African datasets from different sources are:

- Identify the sources and define goals (this includes determining the format, structure, and content of the data, identifying the specific data elements that need to be harmonized and the desired format for the harmonized data) - Given the huge amount of work that has already been carried out in establishing a INSPIRE-compliant dataset for pan-European mineral occurrences and deposits (e.g., Waardenburg, 2014), the Minerals4EU DB structure was used as a template to map the available African data. However, there is also an established DB termed “Sig-Afrique”, which contains minerals data from Africa and is owned by BRGM. Other African data was sourced from individual; in part, mostly confidential, mining and government reports, made available to LNEG in projects undertaken in southern Africa.

- Develop a plan and perform the harmonization (based on the goals and the data sources, develop a plan for how to harmonize the data) - To optimize time and resources, all information is merged to create a single INSPIRE-compliant dataset that has the existing data and will be updated with newer data. This newer data will be gathered from independent reports and commercial industry intelligence data sources (e.g., S&P Capital IQ). This integration includes creating a mapping of the data elements from the different sources and determining any necessary data transformations or cleaning that need to be performed. The data cleaning performed involves identifying and removing errors, inconsistences, duplicate data, correcting spelling mistakes, standardizing units of measurements and removing outliers (Fig. 7).

- Validating the results – After the data is harmonized, it is important to validate the results to ensure that the data is accurate and complete. This include performing quality checks, comparing the normalized data to the original sources, and testing the data to ensure that it can be used effectively.

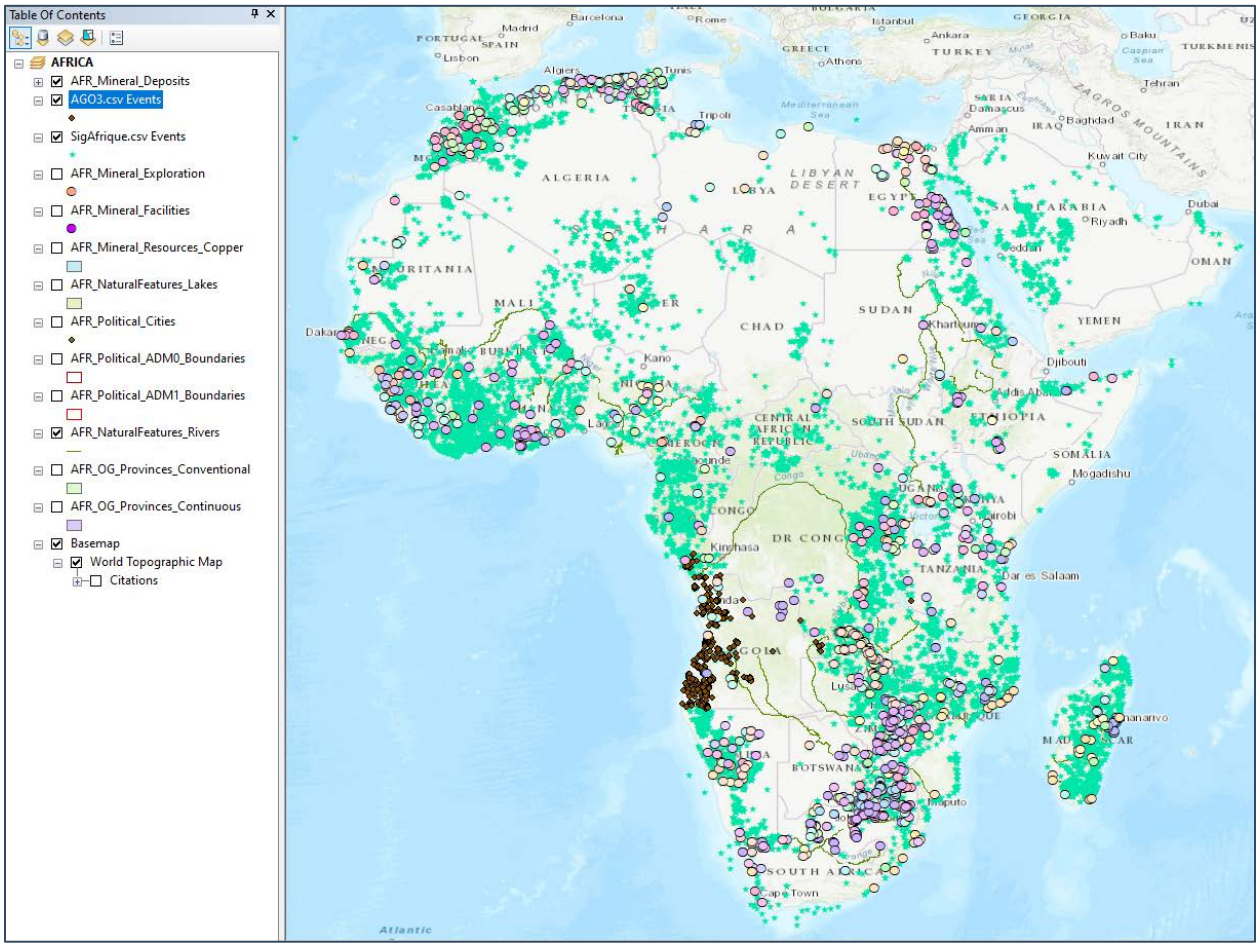


Figure 7 – Overview of some mineral occurrences data sources for Africa: SigAfrique dataset (BRGM), Angola dataset (LNEG), USGS dataset (Padilla et al., 2021).

2.1.3 Data confidentiality

The only data used that does not fall into the confidential category is the one that is obtained freely online through the web. All other data used is to be considered confidential. In order to overcome this limitation, the data needs to be presented in such a way as not to compromise the sources and the confidentiality instilled within them.

Reserves data is excluded for two reasons: 1- it may now be outdated, and 2- it belongs to the various companies, and it may have financial implications in terms of economic reports for shareholders should it come to light.

For this reason, the AfricaMaVal team agreed about the two data format that will make available:

- Web Map Service (WMS) - The WMS specification is a standard that makes it possible to provide geographic information in the form of georeferenced images over the Internet. The WMS for a mineral occurrences map is typically read-only, meaning that it can only be used to view the data specified before. We will make the Mineral predictability maps (using CBA methodology) and others considered important without information on values of reserves, resources, and production of mineral substances, available in this manner.

- Web Feature Service (WFS) – A WFS service is a specification of a format that provides geographic data in vector format, i.e., information is communicated in the form of vectors and attributes that describe them. These Geographic Data Services (GDS) can be used using a GIS program or application (Geographic Information System) compatible with these standards, free (e.g., Quantum GIS, gvSIG, uDig), or commercial (e.g., ArcGIS, Geomedia, AutoCAD Map).

With the WFS format, the dataset can also be used to update and manipulate the data. For example, a user might use a WFS service to query the dataset for mineral occurrences meeting certain criteria, and then use a WMS map to view the locations of the resulting deposits. Here we must specify what kind of data the user can query and see, because the user can have access to the underlying data. In the EGDI platform, which provides access to Pan-European and national geological datasets and services from the Geological Survey Organizations of Europe, a number of European data harmonization projects are accessible, but for mineral resources a user only can query the service about: location, commodities, Mines, Mining activity, Mine status.

2.1.4 Data management and visualization

Managing and visualising the AfricaMaVal data can be undertaken using, for example, a geological data portal such as LNEG's Geoportal.

LNEG's geoPortal (https://geoportal.lneg.pt) is LNEG's integrated services infrastructure to support the management and visualization of spatial data, which aims to provide, in a web environment, the georeferenced information related to the different institutional activities of the National Laboratory of Energy and Geology.

This tool also allows implementing and securing data services (WMS and WFS) in accordance with internationally established technical specifications by organizations related to spatial data.

The GeoPortal consists of four main components:

- Metadata Catalogue: a service for research and consultation of institutional metadata, which allows to know the existence and availability of the geographical information of the LNEG;
- Online Databases: a set of applications that allow access and consultation of information on various institutional themes;
- Map Viewer: LNEG mapping and spatial information visualization and analysis service, which also allows the overlay and spatial analysis of geographic data from different sources (maps, geoscientific data, orthophotomaps, etc.).
- Cartography Download: Space for the provision of cartography following the open data policy.

These four components allow the user to search for spatial information made available by the Institution, to consult and analyze them (as queries or map services) or to download official maps.

3 Overview of the final database

The final AfricaMaVal_GIS geodatabase will contain the following supporting data:

- Data in **shapefile (.shp)** format for each feature class contained in the geodatabase, which may be used to independently load the data into any GIS application capable of reading geospatial data.
- **Layer file (.lyr)** format for each feature class contained in the geodatabase, which may be used to load the symbology and formatting used by the authors in the original geodatabase.
- Data in **layer package (.lpk)** format for each feature class contained in the geodatabase, which includes all field properties and symbology for each respective feature class embedded within the package; note that these files are only compatible with ArcGIS Pro, ArcGIS 10, and ArcGIS 9.3.1 software.

- A **reference table** in excel spreadsheet (.xlsx) format with the attribute properties for all feature classes contained in the geodatabase; the table includes the field aliases and descriptions (which are not contained in the shapefiles) for all fields in each feature class of the geodatabase; this reference table is provided to aid with interpretation of the data if used independently from the geodatabase file and its associated metadata.

The final geodatabase, in its version 1.0, will comprises the following attributes in a reference table (Table 4):

Table 4 - The complete attributes reference table for AfricaMaVal database version 1.0

FIELD NAME	ALIAS	DESCRIPTION	DATA TYPE	Number Format
FID	FID	Internal feature number	Object ID	
Shape	Shape	Feature geometry	Geometry	
FeatureUID	Deposit ID	Unique identification number (ID) given to each deposit in this dataset; IDs are composed of the 3-character alpha GENC country code combined with a sequentially assigned number per deposit per country. Used as labels on map	Text	
Latitude	DD Latitude	Latitude in decimal degrees	Double	Numeric
Longitude	DD Longitude	Longitude in decimal degrees	Double	Numeric
Country	Country (Short Form)	Short form name of the country in which the feature is located, listed in accordance with the Country Codes Type (www.iban.com/country-codes). Alpha-3 code	Text	
OccurrenceName	Occurrence Name	The name of the mineral occurrence/deposit	Text	
OccurrenceOtherName	Occurrence Name_other	Other names for the Occurrence/deposit	Text	
OccurrenceType	MineralOccurrenceType	The type of mineral occurrence. Eg: prospect, occurrence, mineral deposit, ore deposit. See INSPIRE MineralOccurrenceType_CodeList	Text	
DepositGroup	MineralDepositGrouptype	A grouping of mineral deposits defined by generic characteristics. See INSPIRE DepositGroupType_CodeList	Text	
DepositType1	MineralDepositType_1	Style of mineral occurrence or deposit. See INSPIRE DepositType_CodeList	Text	
DepositType2	MineralDepositType_2	Style of mineral occurrence or deposit. See INSPIRE DepositType_CodeList	Text	
HostRock1	RockMaterialLithologyType_1	A controlled concept indicating the name of the RockMaterial type (eg, quartz sandstone, basalt, muscovite schist, sand, mud, soil, saprolite). See INSPIRE LithologyType_CodeList	Text	
HostRock2	RockMaterialLithologyType_2	A controlled concept indicating the name of the RockMaterial type (eg, quartz sandstone, basalt, muscovite schist, sand, mud, soil, saprolite). See INSPIRE LithologyType CodeList	Text	
CRM_1	Critical Raw Material Type_principal	Primary commodity of interest in the deposit, as listed in the source study (See INSPIRE CommodityType_Codelist)	Text	
CRM_2	Critical Raw Material Type_others	Secondary commodity(es) of interest in the deposit, as listed in the source study (See INSPIRE CommodityType_Codelist)	Text	
CRM_Subst	Critical Raw Material Type_Substance	Critical Raw Materials for which we do have the figures for resources, reservces and production. See INSPIRE CommodityType_Codelist	Text	
OtherCommodities	Other commodities	Other commodities not belonging to the CRM list. See INSPIRE CommodityType_Codelist	Text	

DepositSize	Deposit Size_Critical Raw Materials	The importance of the deposit for the commodity. Ex: Occurrence, small deposit, medium deposit. See INSPIRE ImportanceType_Codelist	Text	
OreMineral1	Ore Mineral Name Type_1	Name of the mineral (eg: orthoclase) or mineral family (eg: feldspar), approved by the International Mineralogical Association. (eg: http://www.mindat.org/mineralindex.php). See INSPIRE MineralNameType_Codelist.	Text	
OreMineral2	Ore Mineral Name Type_2	See INSPIRE MineralNameType_Codelist	Text	
OreMineral3	Ore Mineral Name Type_3	See INSPIRE MineralNameType_Codelist	Text	
Gangue1	Gangue Mineral Name Type_1	See INSPIRE MineralNameType_Codelist	Text	
Gangue2	Gangue Mineral Name Type_2	See INSPIRE MineralNameType_Codelist.	Text	
AlterationType1	Alteration Description Type_1	Alteration Distribution describes the spatial distribution or geometry of alteration zones. See INSPIRE AlterationType_Codelist	Text	
AlterationType2	Alteration Description Type_2	See INSPIRE AlterationType_Codelist	Text	
Shape	Occurrence Shape type	The typical geometrical shape of the Earth Resource. See INSPIRE ShapeType_Codelist	Text	
Form	Occurrence Form type	The orebody's typical physical and structural relationship to wallrocks and associated rocks. See INSPIRE FormType_Codelist	Text	
MineStatus	Mine Status Type	Operational status value of the mine. See INSPIRE MinestatusType_Codelist	Text	
MiningActivity	Mining Activity Type	The type of mining activity. Eg: Open Pit, Underground Mine, ... See INSPIRE MiningActivityType_Codelist	Text	
ResourceCategory	Resource Category Type_1	Indication of whether the resource is measured, indicated, or inferred. See INSPIRE ResourceCategoryType_Codelist	Text	
Volume	Volume	The volume of the commodity	Text	
Grade	Grade	The grade of the commodity	Text	
CutoffGrade	Cut off Grade	The cutoff grade used for calculating the commodity measure	Text	
ResourceDate	Resource_Category date of issue	Reference base year of data source material	Long	Numeric
ReserveCategory	Reserves Category Type_1	The level of confidence of the estimate (proved, probable). See INSPIRE ReserveCategoryType_Codelist	Text	
Volume	Volume	The volume of the commodity	Text	
Grade	Grade	The grade of the commodity	Text	
ReserveDate	Reserve Category date of issue	Reference base year of data source material	Long	Numeric
ClassificationMethod	Ore Measure classification Method Used	Means of calculating the measurement. See INSPIRE ClassificationMethodUsedType_Codelist	Text	
ProductionVol	Production Volume	The volume of the commodity	Text	
ProductionGrade	Production Grade	The grade of the commodity	Text	
ProductionStart	Production date_Start	The year production began	Long	Numeric

ProductionEnd	Production date_End	The year production ended	Long	Numeric
DataProvider	Data Provider Contact	Name of organization - Providing data	Text	
Date	Date - Data Provider	Date and time at which this version of the spatial object was inserted or changed in the spatial data set.	Long	Numeric
Reference	Reference	Name of the document. Source and/or reference ID for the source material(s) used to determine the attribute information of mineral deposits.	Text	
Comments	Comments	Any additional noteworthy comments or observations	Text	

The final AfricaMaVal web services will disseminated via:

- **WMS** - List of Web Map Services (WMS) that allow the visualization of maps dynamically. The addresses made available must be consumed in a GIS application compatible with the integration of WMS services. The WMS of the AfricaMaVal DB can be found at the following address:

  https://sig.lneg.pt/server/services/Projects/AfricaMaVal/MapServer/WMSServer?

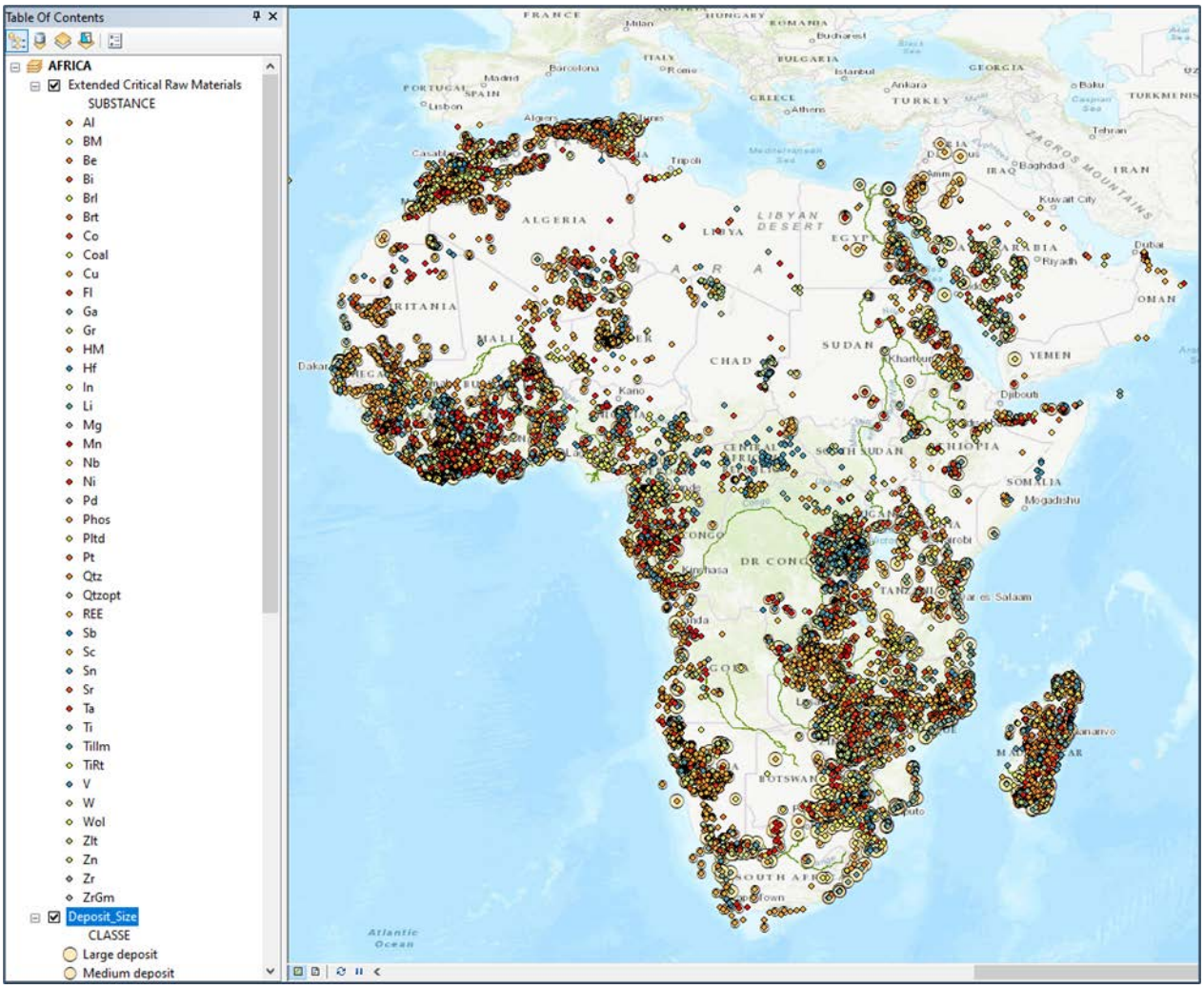


Figure 8 – Overview of AfricaMaVal GIS geodatabase showing the ECRM distribution and their Important Type (Size of the occurrence/deposit).

Overall, we can estimate, at this point, that the AfricaMaVal GIS geodatabase will comprise in this early stage of development more than 20 000 mineral occurrences. However, before the final AfricaMaVal GIS geodatabase is terminated and as further data is added, it will have more than 22 000 data points.

4. Conclusion

For ease of management, the AfricaMaVal GIS geodatabase started out as a hybrid INSPIRE-SigAfrique DB due to the incompatibility between INSPIRE-Compliant and non-INSPIRE-Compliant data.

The data then underwent a series of processes, namely:

- Identified the sources and define the goals - determining the format, structure, and content of the data, identifying the specific data elements that need to be harmonized and the desired format for the harmonized data;
- Developed a plan and perform the harmonization (based on the above point) - To optimize time and resources, all information was merged to create a single INSPIRE-compliant dataset that has the existing data and will be updated with newer data. The data cleaning performed involves identifying and removing errors, inconsistences, duplicate data, correcting spelling mistakes, standardizing units of measurements and removing outliers; and, finally,
- Validating the results by performing quality checks, comparing the normalized data to the original sources, and testing the data to ensure that it can be used effectively.

Bibliography

Cassard, D., Tertre, F., Bertrand, G., Tellez-Arenas, A., Schjøth, F., Tulstrup, J., Heijboer, T., Vuollo, J., Cápová, D., Šinigoj, J., Gruijters, S., Tomas, R. & Schubert, C., (2014a). The Minerals4EU Knowledge Data Platform for managing Web-based mineral resources information in Europe. Minerals in Circular Economy Conference (MINCE). Espoo, Finland, 26−27 November 2014.

Cassard, D., Tertre, F., Tulstrup, J., Vuollo, J., Tomas, R., Capova, D., Sinigoj, J., Delfini C & Burlet, C., (2014b). The Minerals4EU Knowledge Data Platform: implementing INSPIRE data models in the first pan-european raw materials knowledge base. INSPIRE Conference 2014. Aalborg, Denmark, 16-20 June 2014.

Cassard, D., Tertre, F., Bertrand, G., Tellez-Arenas, A., Schjøth, F., Tulstrup J., Heijboer, T., Vuollo, J., Čápová, D., Šinigoj, J., Požar, M., Gruijters, S., Bogaard, P., Waardenburg, F., Tomas, R. & Schubert, C., (2015). Minerals4EU - INSPIRE in Action: Pan-European Raw Materials Knowledge Base Data Platform. INSPIRE Conference 2015. Lisbon, Portugal, 25−29 May 2015.

Fortes, C., Viegas, H, Oliveira, D., Filipe, A., Almeida, P & Gruijters S., (2013). EUROGEOSOURCE - the new generation of EU mineral and energetic resources Web GIS Systems using cloudcomputing. In: Geonovas, no 26, p.3-10. EuroGeoSource Project, http://www.eurogeosource.eu/

FRAME Project, 2018. Forecasting and Assessing Europe's Strategic Raw Materials needs. Deriverable: D3.1: Producing a report describing the methodology used for the identification and selection process of the CRM to be included in the metallogenetic map. Work Package: WP3. http://www.frame.lneg.pt

INSPIRE, 2013. INSPIRE Thematic Working Group Mineral Resources, D2.8.III.21–INSPIRE Data Specification on *Mineral Resources* –Technical Guidelines, version 3.0, pp.156. https://INSPIRE.ec.europa.eu/

Kumar, G.; Basri, S.; Imam, A.A.; Khowaja, S.A.; Capretz L.F.; Balogun, A.O., (2021). Data Harmonization for Heterogeneous Datasets: A Systematic Literature Review. Appl. Sci. 2021, 11, 8275. https://doi.org/10.3390/app11178275

Minerals4EU Project, 2014. European Mineral Knowledge Data Platform, http://www.minerals4eu.eu/, http://minerals4eu.brgm-rec.fr/

Padilla, A.J., Otarod, D., Deloach-Overton, S.W., Kemna, R.F., Freeman, P.A., Wolfe, E.R., Bird, L.R., Gulley, A.L., Trippi, M.H., Dicken, C.L., Hammarstrom, J.M., and Brioche, A.S., (2021). Supporting Data for USGS data release entitled: "*Compilation of Geospatial Data (GIS) for the Mineral Industries and Related Infrastructure of Africa*". [DOI: https://doi.org/10.5066/P97EQWXP]

Schwabe, Craig & Govender, Sives, (2009). Getting Geoinformation and Spatial data Infrastructure (SDI) to work for Africa: The rationale behind Aesi-Align. https://www.researchgate.net/publication/237324083

Vuollo, J., Eloranta, T., Cassard, D., Tomas, R. & Schubert, C., (2014). Geoscience Data Transfers Standards: EarthResourceML and GeoSciML, tools to deliver mineral resources data in EU and globally. The EU-MKDP example. Minerals in Circular Economy Conference (MINCE). Espoo, Finland, 26–27 November (2014).

Waardenburg, Frank D.E. (TNO), Heijboer, Tjerk C. (GEUS), Pen, Simon (TNO), Schjøth, Frands (GEUS) and WP5 Team, (2014). Minerals4EU, WP5. The Enterprise Architect UML and model for M4EU Database.

1. Annex 1 - Minerals4EU-WP5: The conceptual model for M4EU Database

Minerals4EU FP7-NMP.2013.4.1-3

Co-funded by the European Union

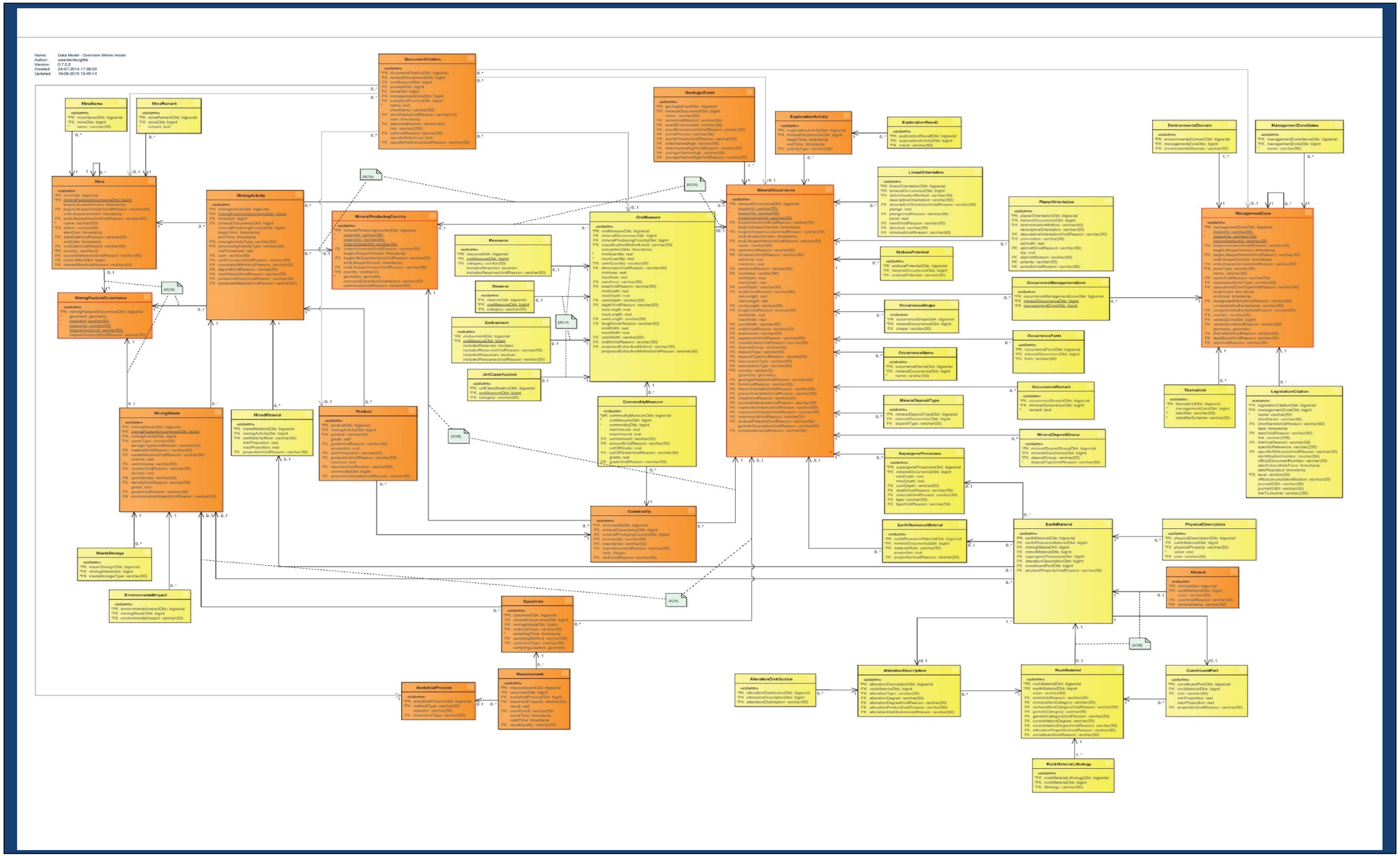
Minerals Intelligence Network for Europe

Minerals4EU-WP5:

The Enterprise Architect UML and model for M4EU Database

Title of the project:	Minerals Intelligence Network for Europe – Minerals4EU
Grant Agreement number:	608921
Funding Scheme:	FP7-NMP-2013: 4.1-3 –CSA (COORDINATING)
Start date:	01.09.2013
Duration:	24 months
Document title:	Minerals4EU-WP5: The Enterprise Architect UML and model for M4EU Database
Workpackage:	WP5
Author(s):	Frank D.E. Waardenburg (TNO), Tjerk C. Heijboer (GEUS), Simon Pen (TNO), Frands Schjøth (GEUS) and WP5 Team
Date of delivery:	20.08.2014
Dissemination level:	**PU**/PP/RE/CO[1]
Reviewed by:	
Status of the document:	**Final**
Document location:	Folder: Minerals4EU/WP5/M4EU data model
Project web site:	http://www.minerals4eu.eu

[1] PU=public, PP= Restricted to other programme participants (including the Commission Services), RE= Restricted to a group specified by the consortium (including the Commission Services), CO= Confidential, only for members of the consortium (including the Commission Services)


Name: Data Model - Overview Whole model
Author: waardenburgfde
Version: 0.7.0.2
Created: 24-07-2014 17:39:33
Updated: 19-09-2015 13:49:14
MineName
MineRemark
Mine
MiningActivity
DocumentCitation
GeologicEvent
ExplorationActivity
ExplorationResult
MineralOccurrence
LinearOrientation
PlanarOrientation
EndusePotential
EnvironmentalDomain
ManagementZoneName
ManagementZone
OccurrenceManagementZone
OccurrenceShape
OccurrenceName
OccurrenceForm
OccurrenceRemark
MineralDepositType
MineralDepositGroup
SupergeneProcesses
EarthResourceMaterial
EarthMaterial
PhysicalDescription
Mineral
ThematicId
LegislationCitation
MineralProducingCountry
Resource
Reserve
Endowment
UnfClassification
OreMeasure
CommodityMeasure
Commodity
MiningFeatureOccurrence
MiningWaste
MinedMaterial
Product
WasteStorage
EnvironmentalImpact
Specimen
Measurement
AnalyticalProcess
AlterationDistribution
AlterationDescription
RockMaterial
ConstituentPart
RockMaterialLithology
(XOR)
(XOR)
(XOR)
(XOR)
(XOR)