# CRMs DataSpace

## RFCS PROJECT

Common European Data Space on Critical Raw Materials for the Green Deal: Recovery Potential from Closed Extractive Coal Waste Facilities

Grant Agreement 101216677

## Deliverable 1.3

Comprehensive Overview of the Project

Co-funded by the European Union

Research Fund for Coal & Steel

## Authors

Prof. Alicja Krzemień, CENTRAL MINING INSTITUTE – NATIONAL RESEARCH INSTITUTE (Poland)

Dr. Antonio Luis Marqués Sierra, UNIVERSITY OF OVIEDO (Spain)

| Deliverable 1.3 | |
|---|---|
| Due date of Deliverable | 31.12.2025 |
| Start - End Date of Project | 01.07.2025 – 30.06.2028 |
| Duration | 3 years |
| Deliverable Lead Partner | GIG-PIB |
| Dissemination level | Public |
| Work Package | WP 1 |
| Digital File Name | D1.3 Comprehensive Overview of the Project |
| Keywords | Project overview, state of the art, approach, outcome |

## Disclaimer

The information and photographs in this Deliverable remain the property of the CRMsDataSpace Project or its Partners. You must not distribute, copy or print this information.

# Table of contents

EXECUTIVE SUMMARY 8

1 INTRODUCTION 10

2 OBJECTIVES 12

2.1 KEY PERFORMANCE INDICATORS (KPI) 14

3 APPROACH 16

3.1 IMPLEMENTING AI REPORTING TOOLS USING LARGE LANGUAGE MODELS (LLMS) 18
3.2 IMPLEMENTING AN AI TOOL TO CLASSIFY CRMS FROM CLOSED EXTRACTIVE WASTE FACILITIES ACCORDING TO THE UNFCR 19

4 METHODOLOGY 21

4.1 CHALLENGE 1: TO DEVELOP THE FIRST DATA CATALOGUE THAT LEVERAGES METADATA AND DATA MANAGEMENT TOOLS FOCUSED ON CRITICAL RAW MATERIALS 22
4.2 CHALLENGE 2: TO ENSURE DATA EXCHANGE, INTEGRATION, AND ANALYSIS BETWEEN CRM STAKEHOLDERS BY DEVELOPING AND DEPLOYING A DATA SPACE PLATFORM ALIGNED WITH THE PRINCIPLES GOVERNING EUROPEAN DATA SPACES 23
4.3 CHALLENGE 3: TO DEVELOP MULTI-SIDED BUSINESS SOLUTIONS FOR THE CRM ECOSYSTEM 26
4.4 CHALLENGE 4: TO IMPLEMENT ARTIFICIAL INTELLIGENCE (AI) REPORTING TOOLS USING LARGE LANGUAGE MODELS (LLMS) 27
4.5 CHALLENGE 5: TO IMPLEMENT AN ARTIFICIAL INTELLIGENCE TOOL TO CLASSIFY CRMS FROM CLOSED EXTRACTIVE WASTE FACILITIES ACCORDING TO THE UNITED NATIONS FRAMEWORK CLASSIFICATION FOR RESOURCES 29
4.6 CHALLENGE 6: TO DEPLOY AN OPERATIONAL COMMON EUROPEAN DATA SPACE ON CRITICAL RAW MATERIALS FOR THE GREEN DEAL SUPPORTED BY CRMSDATASPACE 30
4.7 WORK PLAN OF THE PROJECT 30

5 IMPACT 32

5.1 RELEVANT AND SUBSTANTIAL EXPECTED BENEFITS FOR THE EUROPEAN COAL SECTOR 32
5.2 SCALE AND SIGNIFICANCE OF THE PROJECT'S CONTRIBUTION TO THE EXPECTED OUTCOMES AND IMPACTS IN THE EUROPEAN COAL SECTOR 33
5.3 REQUIREMENTS AND POTENTIAL BARRIERS 34
5.4 TARGET GROUPS, POTENTIAL USERS, STAKEHOLDERS, AND MULTIPLIERS 34

6 ANTICIPATED BENEFITS 35

7 GLOSSARY 37

8 ANNEX: STATE OF THE ART 38

## List of Figures

Figure 3-1. UNFCR Categories and Examples of Classes ..... 20
Figure 4-1. The overall methodology of CRMsDataSpace ..... 21
Figure 4-2. Block model and deposit modelling for resource estimation ..... 23
Figure 4-3. How the UNFCR works ..... 29
Figure 4-4. Work plan of the CRMSDataSpace project ..... 31
Figure 8-1. EMODnet Marine Critical Minerals Map Viewer ..... 40
Figure 8-2. Geoportal of the Federal Institute for Geosciences and Natural Resources 41
Figure 8-3. IGME Spanish National Inventory of Mining Tailings ..... 41
Figure 8-4. Geological Survey Ireland Spatial Resources ..... 42
Figure 8-5. OPI-TPP 2.0 ..... 43

## List of Tables

Table 2-1. Key Performance Indicators. ..... 14
Table 3-1. Common European Data Space ecosystems and sub-ecosystems. ..... 17

## Executive Summary

CRMsDataSpace is an ambitious initiative to establish the first Common European Data Space on Critical Raw Materials (CRMs) for the Green Deal, focusing on the recovery potential of CRMs from closed extractive coal waste facilities across Europe. The project responds to the objectives of the Critical Raw Materials Act (CRMA) and contributes to strengthening the European Union's strategic autonomy, industrial competitiveness, and green transition.

The project combines expertise from mining sciences, digital technologies, artificial intelligence, and data governance to create a secure, interoperable, and trustworthy ecosystem for sharing and exploiting critical raw materials data. By deploying a Minimum Viable Data Space (MVDS), CRMsDataSpace will facilitate data exchange among stakeholders, including mining companies, public authorities, geological surveys, research organisations, and technology providers.

A key innovation of the project is the development of the first dedicated data catalogue focused on critical raw materials from closed extractive waste facilities. The platform will integrate geological, geochemical, mineralogical, environmental, spatial, and operational data, enabling efficient discovery, sharing, and reuse of information throughout the CRM value chain.

CRMsDataSpace will also pioneer the application of Artificial Intelligence (AI) and Large Language Models (LLMs) within the critical raw materials sector. AI-powered reporting tools will automatically generate thematic reports to help project promoters, policymakers, investors, and researchers identify recovery opportunities, assess economic viability, and explore strategic supply chain options. In parallel, a novel AI-based classification tool will support the categorisation of CRM projects in accordance with the United Nations Framework Classification for Resources (UNFC), thereby facilitating compliance with emerging European regulatory requirements.

The project will develop multi-sided business solutions that create value for both data providers and data users. These solutions include enablement services such as GIS-based visualisation and data-sharing tools, as well as business applications for field data collection, georeferenced information management, and CRM resource evaluation. By leveraging network effects, the ecosystem is expected to attract increasing numbers of participants, datasets, and services over time.

To demonstrate the practical applicability of the platform, CRMsDataSpace will populate the ecosystem with case studies from multiple European countries, integrating primary data generated through sampling, characterisation, resource estimation, and recovery studies, as well as secondary data from relevant EU-funded projects.

The expected impacts of CRMsDataSpace extend beyond technological innovation. The project will contribute to the identification and recovery of critical raw materials from abandoned mining waste facilities, support the rehabilitation of degraded mining areas, stimulate economic activity and job creation in coal regions in transition, and enhance the resilience of European critical raw materials supply chains. Furthermore, it will strengthen the development of the European data economy by establishing a dedicated CRM ecosystem aligned with the principles of Common European Data Spaces.

Through its combination of data space technologies, artificial intelligence, resource assessment methodologies, and stakeholder engagement activities, CRMsDataSpace will create the foundations for a scalable and sustainable European ecosystem capable of accelerating the recovery, management, and strategic use of critical raw materials in support of the Green Deal and Europe's industrial future.

# 1 Introduction

This deliverable (D1.3) provides a comprehensive overview of the CRMsDataSpace project, an initiative funded under the Research Fund for Coal and Steel (RFCS) that aims to establish the first Common European Data Space on Critical Raw Materials (CRMs) for the Green Deal. The project seeks to support the implementation of the Critical Raw Materials Act (CRMA) by creating a secure, interoperable and trustworthy ecosystem for sharing and exploiting data related to critical raw materials, particularly those contained in closed extractive coal waste facilities across Europe.

CRMsDataSpace focuses on developing a Minimum Viable Data Space (MVDS) that enables the integration, exchange, and analysis of data among stakeholders in the critical raw materials value chain, including public authorities, geological surveys, mining companies, research organisations, and technology providers. Through this approach, the project aims to improve access to critical raw materials information, facilitate informed decision-making and strengthen the resilience of European raw materials supply chains.

The project leverages synergies between mining sciences, information technologies, artificial intelligence, data governance frameworks and stakeholder engagement activities. This interdisciplinary approach allows CRMsDataSpace to address challenges related to data accessibility, interoperability, regulatory compliance and the identification of critical raw materials recovery opportunities from closed extractive waste facilities.

Several innovative components will be developed within the project, including the first dedicated data catalogue focused on critical raw materials from closed extractive waste facilities, a data space platform aligned with European Data Space principles, multi-side business solutions for the CRM ecosystem, Artificial Intelligence reporting tools based on Large Language Models (LLMs), and an AI-based tool for the classification of critical raw materials according to the United Nations Framework Classification for Resources (UNFC). These developments will contribute to the deployment of an operational Common European Data Space on Critical Raw Materials for the Green Deal.

This deliverable is part of Work Package 1 and is entitled “D1.3 Comprehensive Overview of the Project”. The purpose of this deliverable is to provide a clear and structured description of the project, including its objectives, methodology, expected outcomes, key performance indicators, stakeholder engagement approach and anticipated impacts.

In particular, Deliverable D1.3 plays a transversal role by presenting the overall vision of CRMsDataSpace and establishing a common understanding of the project among partners, stakeholders and external audiences. It serves not only as a formal project

overview, but also as a communication tool that highlights the project's contribution to the European Green Deal, the Critical Raw Materials Act and the development of a data-driven ecosystem capable of supporting the sustainable management and recovery of critical raw materials in Europe.

# 2 Objectives

Critical raw materials (CRMs) are non-energy, non-agricultural raw materials that play a vital role in the EU economy and are at high risk of supply shortage. These materials are essential for various strategic sectors, including renewable energy, digital industry, space and defence, and healthcare. For instance, lithium, cobalt, and nickel are used in battery production, gallium in solar panels, and raw boron in wind technologies.

The European Critical Raw Materials Act (CRMA)[1] was enforced on 23 May 2024 to ensure a sustainable and secure supply of critical raw materials for Europe's industry. The EU aims to significantly lower its dependency on imports from single-country suppliers. This legislation has quickly become a landmark policy to secure sustainable supply chains for critical minerals.

The EU is again at the forefront, setting ambitious targets for ethical sourcing and transparency. The CRMA is not just about regulation; it's a call to action, aiming to reduce the environmental and social impacts of mineral extraction and processing while ensuring these resources are managed for future generations.

The CRMA identifies a list of critical and strategic raw materials essential for technologies used in the green and digital transition, defence, and space. It also establishes targets for domestic capacities along the supply chain to achieve these strategic raw materials by 2030. At any relevant stage of processing, no more than 65% of the EU's annual needs for each strategic raw material should come from a single third country.

One of CRMA's main goals is to create a more sustainable and circular economy for CRMs. This includes promoting the recycling of raw materials and supporting a robust secondary market. To achieve this, Member States will be required to intensify their efforts to recover CRMs from waste products and mining waste. Additionally, there will be encouragement to recover CRMs from closed extractive waste facilities.

Member States should promote the recovery of critical raw materials from extractive waste by improving the availability of information and by addressing legal, economic and technical barriers. To address the current lack of information on the potential of closed extractive waste facilities for critical raw materials, the CRMA requires Member States to establish a database that contains all relevant information to promote

[1] European Critical Raw Materials Act. Regulation (EU) 2024/1252 of the European Parliament and of the Council of 11 April 2024 establishing a framework for ensuring a secure and sustainable supply of critical raw materials and amending Regulations (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1724 and (EU) 2019/1020.

recovery, particularly the quantities and concentrations of critical raw materials in such facilities.

These databases shall contain all closed waste facilities on the territory of each Member State, including abandoned ones, except for those where the particular characteristics of the waste sites or geological conditions make the presence of potentially technically recoverable quantities of CRMs unlikely.

Many Member States are expected to use existing databases created by their national geological services, often for other purposes, to meet the CRMA's requirements. European-level databases have also been developed under various European projects; this will cause severe dysfunction when using tools developed with different objectives and languages across Member States, and will prevent comparisons at the European level.

Building a Common European Data Space for critical raw materials is the optimal way to meet the information requirements imposed by the CRMA while giving them a real European dimension. Data exchange across organisations or Member States is currently constrained by proprietary, non-transparent, non-interoperable technologies that do not provide the necessary level of trust.

The Specific SMART objectives defined for the CRMsDataSpace are as follows:

1. To provide National Authorities in Member States with a tool to comply with the requirements of the CRMA regarding information about CRM's potential in closed extractive waste facilities.
2. To develop multi-sided business solutions for the CRM ecosystem that leverage network effects, catering to both the "supply" and "demand" of data and creating data-related services, including enablement services to provide or consume data and business applications to solve specific problems within the ecosystem.
3. To implement Artificial Intelligence (AI) reporting tools using Large Language Models to generate reports automatically of great interest to potential project promoters, facilitating the search and gathering of information on CRMs.
4. To feed the ecosystem with information from closed extractive coal waste facilities to exemplify assessing the critical raw materials recovery potential by multiple options regarding general information, chemical and mineralogical characterisation from geochemical sampling, mineralurgical and hydrometallurgical processes, reserve estimation based on core logging when relevant, and resource classification according to the UN Framework Classification for Resources.
5. To use secondary data (data not generated by the data space) to feed the Artificial Intelligence reporting tools from relevant EU-funded projects.

6. To develop international workshops to attract the interest of other operators and Member States to provide information to the data space. A workshop will be organised in all member states represented in the consortium: Poland, Spain and Romania. The workshops will be open to authorities from neighbouring countries and regions. A final workshop in Brussels will bring project outcomes near all Member States (through the industry officers of the Permanent Representations), regional and local authorities (through the Committee of the Regions), and other stakeholders (through the European Parliament and the European Economic and Social Committee).
7. To develop specific webinars that address data users and consumers to reach a large number of stakeholders trained and skilled in using the data space.

## 2.1 Key performance indicators (KPI)

In line with the above specific objectives, the Key Performance Indicators are presented in Table 2-1. All the objectives are realistic and achievable during the CRMSDataSpace project duration.

**Table 2-1. Key Performance Indicators.**

| # | KPI description | Related objective | KPI range | Related task/date |
|---|---|---|---|---|
| **1** | Number of national authorities in charge of elaborating and implementing the CRMA Regulation with which direct contact will be established. | Objective 1: To provide National Authorities in Member States with a tool to comply with the requirements of the CRMA. | 27 (all Member States). | Task 7.2 Completion date: Month 36 |
| **2** | Number of multi-side business solutions implemented for the ecosystem. | Objective 2: To develop multi-side business solutions that leverage network effects, catering to both the supply and demand of data and creating data-related services. | At least 5 multi-side business solutions covering different needs of the critical raw materials ecosystem. | Tasks 5.2, 5.3, 5.4, 6.2 and 6.3 Completion date: Month 33 |
| **3** | Number of report templates based on Artificial Intelligence (AI). | Objective 3: To implement Artificial Intelligence (AI) reporting tools using Large Language Models to generate reports automatically. | At least 4 report templates implemented using 2 different base Large Language Models. | Task 6.2 Completion date: Month 27 |

| | | | | |
|---|---|---|---|---|
| **4** | Number of case studies used to feed and exemplify the Critical Raw Materials ecosystem functioning. | Objective 4: To feed the ecosystem with information from closed extractive coal waste facilities to assess the CRM recovery potential in extractive coal mining wastes through multiple options. | At least 4 case studies from 3 different European countries, including resource classification according to the United Nations Framework Classification for Resources using AI. | Task 2.6 Completion date: Month 21 |
| **5** | Number of secondary datasets used to feed the report templates from EU-funded CRM-related projects. | Objective 5: To use secondary data (data not generated by the data space) to feed the Artificial Intelligence reporting tools. | At least 3 datasets from different EU-funded research projects. | Task 6.2 Completion date: Month 27 |
| **6** | Number of international workshops to attract the interest of other operators and Member States to provide information to the data space. | Objective 6: To develop international workshops to attract the interest of other operators and Member States to provide information to the data space. | At least 4 workshops: 3 in different Member States (Poland, Spain and Romania) and 1 in Brussels at the European Parliament. More than 250 attendees. | Task 7.5 Completion date: Month 36 |
| **7** | Number of webinars developed and number of stakeholders trained and skilled in using the data space. | Objective 7: To develop specific webinars that address data users and consumers. | At least 4 webinars and 200 stakeholders trained. | Task 7.3 Completion date: Month 36 |

By meeting these objectives, CRMSDataSpace positions itself as a pioneering initiative in Europe, not only from a technological standpoint but also as a policy-relevant, socially inclusive model for coal regions seeking a viable and sustainable future beyond carbon-intensive industries.

# 3 Approach

The project is ambitious and goes beyond the state of the art as (1) it will start developing the Common European Data Space on Critical Raw Materials for the Green Deal, currently inexistent; (2) it will implement the first Artificial Intelligence (AI) reporting tools using Large Language Models (LLM) to facilitate the search and gathering of information for the recovery potential of CRMs from closed extractive waste facilities; and (3) it will implement a yet non-existent today AI tool for classifying critical raw materials from closed extractive waste facilities according to the United Nations Framework Classification for Resources.

Data is reshaping the way we produce, consume and live. Data is vital to innovative products and services, from real-time navigation to improved personalised medicine, precision farming, and $CO_2$ emissions reduction.

To harness data's value for the benefit of the European economy and society, the European strategy for data set out the path to creating Common European Data Spaces in several strategic fields: health, agriculture, manufacturing, energy, mobility, financial, public administration, skills, the European Open Science Cloud, and the Green Deal. The EU calls “Common European Data Spaces” all data spaces and data spaces initiatives funded by the EU, each typically focused on addressing the needs of one sector (Agriculture, Energy, Finance, Health, Manufacturing, Green Deal, etc.).

They are ‘common’ because they aim to serve the whole of the EU without restrictions to participation. As such, these data spaces are commonly referred to as “open”. At the same time, the common European data spaces are not intended to be exclusive; other data spaces exist and can emerge in the future, serving the same sectors.

The Common European Data Spaces are helping to unleash the enormous potential of data-driven innovation. They make trustworthy, secure data from across the EU available and exchangeable. EU Businesses, public administrations, and individuals will control the data they generate. At the same time, these data holders will benefit from a safe and reliable framework for sharing their data for innovation.

Common European Data Spaces will enhance the development of new data-driven products and services in the EU, forming the core tissue of an interconnected and competitive European data economy. They are being developed across 14 ecosystems and some sub-ecosystems (Table 3-1).

Building a Common European Data Space for critical raw materials is the optimal way to meet the information requirements imposed by the CRMA while giving them a real European dimension. Data exchange across organisations or Member States is currently constrained by proprietary, non-transparent, non-interoperable technologies that do not provide the necessary level of trust.

Table 3-1. Common European Data Space ecosystems and sub-ecosystems.

| Ecosystem | Sub-ecosystems |
|---|---|
| **Agriculture** | - |
| **Cultural Heritage** | - |
| **Energy** | - |
| **Finance** | - |
| **Green Deal** | Smart cities and communities |
| **Health** | Cancer images; Genomics |
| **Language** | - |
| **Manufacturing** | - |
| **Media** | - |
| **Mobility** | - |
| **Public Administration** | Legal; Public procurement |
| **Research & Innovation** | - |
| **Skills** | Edge-Skills |
| **Tourism** | - |

The strategy is to start building a valuable collection of data resources from closed extractive coal waste facilities and services to attract users not only from the coal industry but from the mining industry in general, with the understanding that a large user base will, in turn, attract additional data resources and services.

CRMsDataSpace will initially focus on recovery potential from closed extractive waste facilities, drawing on information from the coal industry. However, it will simultaneously apply to and expand to all extractive waste facilities.

Thus, the project will position the Research Fund for Coal and Steel (RFCS) as a critical player in the EU's efforts to ensure a sustainable and secure supply of critical raw materials for its industry.

In the future, the Common European Data Space on Critical Raw Materials for the Green Deal should introduce other complementary aspects, such as critical raw materials mineral deposits in the EU, national programmes for general exploration targeted at critical raw materials and carrier minerals of critical raw materials (Article 19 of the CRMA), validated strategic project applications, etc.

CRMsDataSpace aims to establish a critical raw materials ecosystem for the Green Deal Common European Data Space for the first time, whereby data is shared and made available in a trustworthy environment. This aim can be translated into three specific ambitions beyond the state of the art:

1. To ensure data exchange, integration, and analysis between CRM stakeholders by developing and deploying a data space platform aligned with the principles governing European data spaces, facilitating the creation of a Minimum Viable Data Space (MVDS).
2. To develop the first data catalogue that leverages metadata and data management tools focused on CRMs, including characterising the different phenomena, attributes, relations, restrictions, controlled terms, multimedia, and general characteristics of the closed extractive waste facilities.
3. To develop multi-side business solutions for the CRM ecosystem that leverage network effects, catering to both the "supply" and "demand" of data and creating data-related services, including (a) Enablement Services to provide or consume data, such as a GIS web portal addressing enhanced visualisation, search, filtering capacity, and download and, conversely, a user-friendly interface to provide data; and (b) Business applications to solve specific problems within the ecosystem such as an app for mobile devices aimed at optimising, homogenising and guaranteeing the quality of data collection in the field and automating the acquisition of GPS georeferenced data including photographs.

## 3.1 Implementing AI reporting tools using Large Language Models (LLMs)

Large Language Models (LLMs) represent an advanced technology in Artificial Intelligence (AI) and Natural Language Processing (NLP). These models are AI systems designed to understand, interpret, and generate human language fluently and accurately, leveraging deep learning techniques, particularly neural networks.

Their ability to process and learn from large volumes of text data allows them to mimic the human ability to predict and generate text that is coherent and contextually appropriate. The evolution of LLMs has been marked by significant advances in NLP since the introduction of the Transformer architecture in 2017, which enabled models to better understand and represent complex linguistic patterns. Since then, models have been developed that have proven extraordinarily effective in various NLP tasks,

including text generation, machine translation, sentiment analysis, and question-answering systems.

A multi-sided business solution will be developed based on large language models (LLMs), which represent advanced technology in Artificial Intelligence (AI) and Natural Language Processing (NLP). The model will automatically generate reports of interest to the data space users, facilitating the search and gathering of information on CRM's recovery potential from closed extractive waste facilities.

Although these reports will mainly use primary data, data directly obtained from the CRM ecosystem, secondary data from other EU projects will also be considered, as with the SCRREEN European project, where exhaustive fact sheets for the critical raw materials were developed.

These Large Language Models (LLMs) will be used for the first time to generate four different types of reports related to CRMs:

1. To identify and fully describe the facility across Europe, within one or several Member States, or a specific region, presenting the most feasible recovery of targeted CRMs.
2. To identify and fully describe the facility/ies across Europe, within one or several Member States or a region with a high potential for economically viable recovery of targeted CRMs for the supply chain in strategic technologies and sectors such as renewable energy, e-mobility, defence and aerospace.
3. To expand the previous reports using secondary data from other EU projects.
4. To draw up the most feasible programmes for general exploration targeted at CRMs and carrier minerals of critical raw materials within a specific area/region/Member State.

Generating these reports will significantly facilitate companies and individuals interested in recovering CRMs from closed extractive waste facilities, undergoing quick and rigorous information searches, thus promoting CRM recovery from these sites within the EU.

## 3.2 Implementing an AI tool to classify CRMs from closed extractive waste facilities according to the UNFCR

The United Nations Framework Classification for Resources (UNFCR) is a project- and principles-based classification system for defining the environmental, socio-economic viability, and technical feasibility of resource development projects. UNFCR provides a consistent framework to describe the confidence level in future quantities the project produces.

UNFCR is a principles-based system in which the products of a resource project are classified based on the three fundamental criteria of environmental-socio-economic viability (E), technical feasibility (F), and degree of confidence in the estimate (G), using a numerical coding system. Combinations of these criteria create a three-dimensional system (Figure 3-1). Categories (e.g. E1, E2, E3) and, in some cases, sub-categories are defined for each of the three criteria.

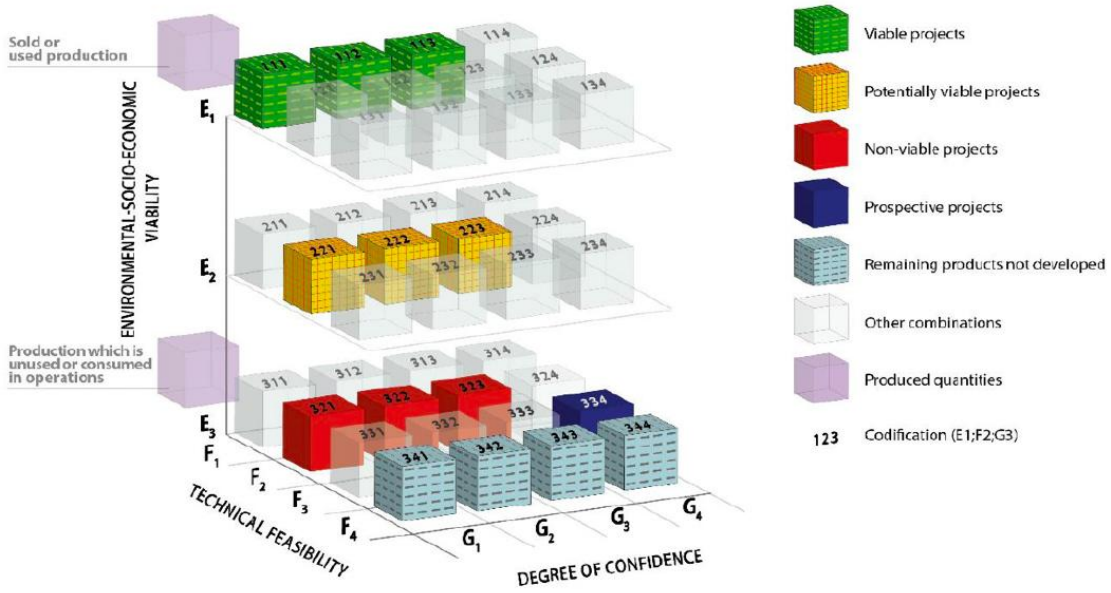



**Figure 3-1. UNFCR Categories and Examples of Classes[2]**

Recital (19) of the CRMA states that any strategic raw materials project promoter should be able to apply to the Commission for recognition of their project as a Strategic Project. The application should include relevant documents and evidence related to the criteria. To better assess the project's social, environmental, and economic viability, feasibility, and confidence level in the estimates, the project promoter should also provide a classification of the project according to the United Nations Framework Classification for Resources.

Artificial Intelligence (AI) will be used for the first time to establish the classification of critical raw materials with high potential for economically viable recovery of critical raw materials according to the *United Nations Framework Classification for Resources*, which has never been done before.

[2] United Nations Framework Classification for Resources - Update 2019. United Nations Economic Commission for Europe. ECE Energy Series No. 61. United Nations. Geneva, 2020.

# 4 Methodology

The following identified challenges, represented in Figure 4-1, correspond to the proposal's different ambitions.

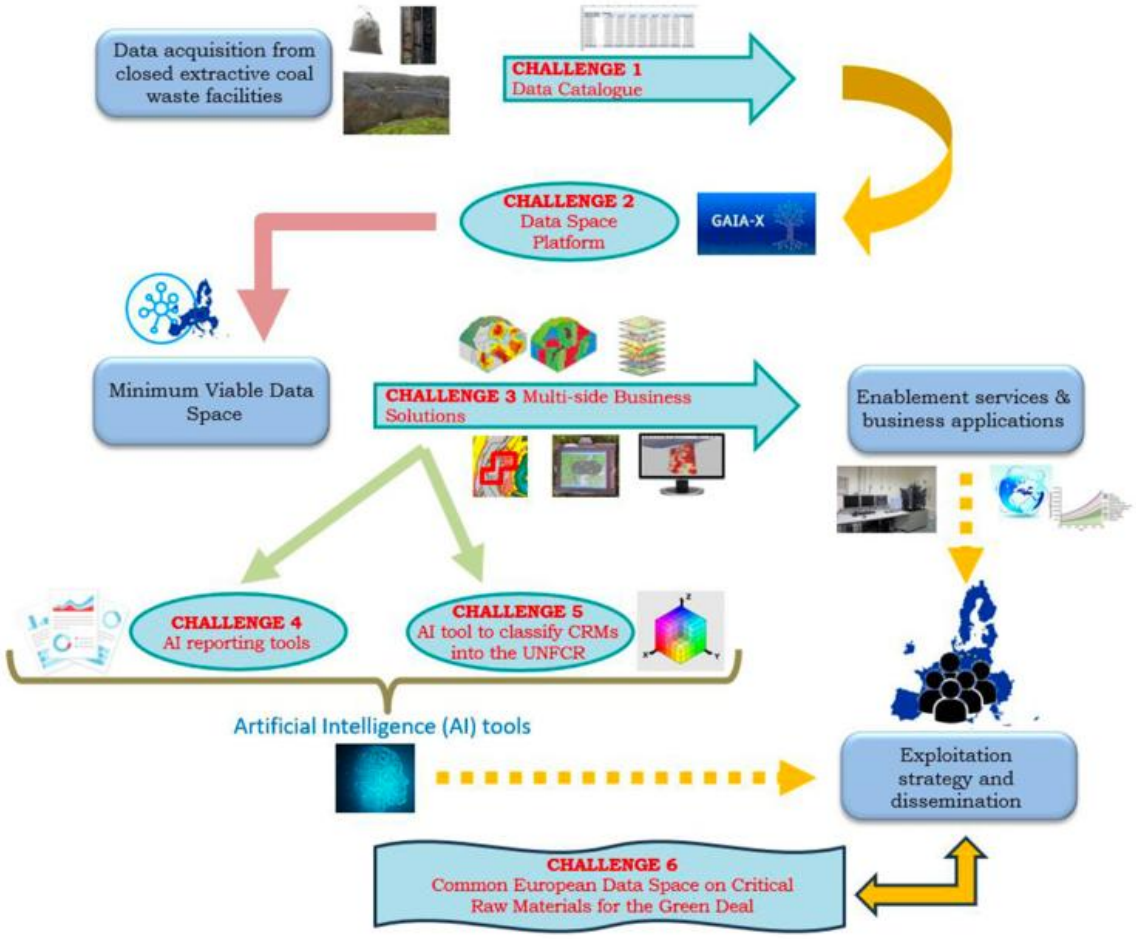



**Figure 4-1. The overall methodology of CRMsDataSpace**

They go beyond the state of the art, using an interdisciplinary approach spanning mining sciences, IT sciences, and social sciences to deploy enablement services and business applications.

- CHALLENGE 1: To develop the first data catalogue that leverages metadata and data management tools focused on critical raw materials.
- CHALLENGE 2: To ensure data exchange, integration, and analysis between CRM stakeholders by developing and deploying a data space platform aligned with the principles governing European data spaces.
- CHALLENGE 3: To develop multi-sided business solutions for the CRM ecosystem.

- CHALLENGE 4: To implement Artificial Intelligence (AI) reporting tools using Large Language Models (LLMs).
- CHALLENGE 5: To implement an Artificial Intelligence (AI) tool to classify CRMs from closed extractive waste facilities according to the United Nations Framework Classification for Resources.
- CHALLENGE 6: To deploy an operational Common European Data Space on Critical Raw Materials for the Green Deal supported by CRMsDataSpace.

## 4.1 CHALLENGE 1: To develop the first data catalogue that leverages metadata and data management tools focused on critical raw materials

It will include characterising the different phenomena with their attributes, relations, restrictions, lists of controlled terms, and multimedia, as well as the general characteristics of the closed extractive waste facilities. To achieve this goal, the following information will be collected in each case study:

1. **Background information:**
    a) Information about the area, history, photographs, nearby populations, geological characteristics, location, spatial coordinates, geometry, 3D representation (based on Google Maps, OpenStreetMap, or similar), digital elevation model, and waste volume (or estimation) of the waste facility.
    b) Where applicable, the waste facility's operator or former operator, legal successor, and administrative status.
    c) The approximate quantities and concentrations of all raw materials contained in the extractive waste and, where available, in the original mineral deposit.
    d) Characterisation of the mining and environmental surroundings.
    e) The mining and beneficiation process for the deposit: extraction, crushing, grinding, separation, and concentration.
    f) The original mineral deposit information, mining legacy data of interest, etc.
2. **Georeferenced geochemical sampling** will be collected in each closed extractive coal waste facility to estimate the quantities and concentrations of critical raw materials. The samples will be selected and distributed according to the facility's area.
3. **Chemical and mineralogical characterisation** of the initial geochemical sampling and the mineralurgical and metallurgical recovery and multihole exploration drilling samples will be developed at accredited laboratories using Inductively Coupled Plasma-Atomic Emission Spectrometers (ICP-AES), X-ray fluorescence (XRF), and Atomic absorption spectroscopy (AAS). Mineral characterisation will

be developed using X-ray diffraction (XRD) and scanning electron microscopy (SEM).

4. **Mineralurgical recovery** will be developed through gravimetric (pulsation box, shaking tables, and multi-gravity concentration), magnetic (magnetic fields and high-field strength magnetic separation), electric (electric fields and high-voltage electric field separation), and froth flotation.
5. **Hydrometallurgical recovery** via lixiviation will be developed using different acids. Critical raw materials will be extracted from the lixiviate liquor using solvent extraction and precipitation. An electrometallurgical test will also be developed for selective recovery of specific critical raw materials.
6. **Multihole exploration drilling** will allow sampling of the soil of the closed extractive coal waste facilities at different depths to determine if a targeted CRM is located along the drilling line. Drilling locates and defines economic CRM and provides the ultimate test for the predictions generated in the preceding geochemical sampling and target generation phases of an exploration process.
7. **Resource estimation** using the RecMin program will allow for the determination and definition of the targeted CRMs' tonnage. The main steps are section plotting, interactive geological modelling, geostatistical analysis, block modelling, and block estimation (Figure 4-2).

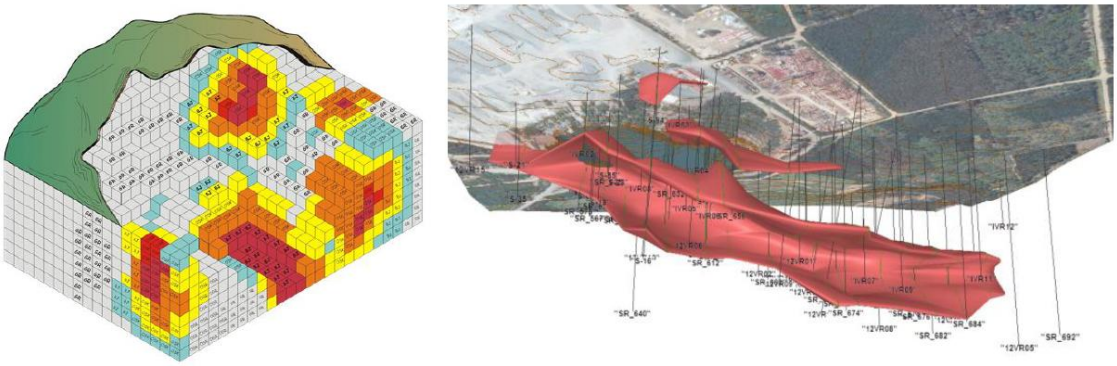


Figure 4-2. Block model and deposit modelling for resource estimation

## 4.2 CHALLENGE 2: To ensure data exchange, integration, and analysis between CRM stakeholders by developing and deploying a data space platform aligned with the principles governing European data spaces

This challenge implies creating a Minimum Viable Data Space (MVDS) and ensuring data exchange, integration, and analysis between stakeholders.

The objectives of this challenge will be to:

- **Design and development**: A scalable, secure, interoperable data space platform that meets the proposal's requirements will be designed and developed.
- **Deployment of the MVDS**: A Minimum Viable Data Space (MVDS) will be deployed on the developed platform, allowing the project's data subject to be catalogued and exchanged with different stakeholders.
- **Definition of the governance framework**: The governance framework that guarantees that the data space platform is managed effectively, efficiently, and following the participants' business objectives will be defined.

To achieve these goals, an agile development methodology will be used. This methodology involves iterative cycles of design improvement, development, testing, and deployment. This approach ensures the platform is developed, incorporating feedback and requirements throughout the project lifecycle.

At the technological level, the proposed architecture by CRMsDataSpace is a hybrid infrastructure based on the following frameworks and protocols, all aligned with European reference frameworks and open source:

**The Eclipse Dataspace Components (EDC)**: An implementation framework of the DSP that enables sovereign data exchange between organisations.

**GAIA-X**: It is a reference model with several layers (technological, governance, organisational, and legal) and is used as the primary building block of the interoperability node or Gaia-X Digital Clearing House (GXDCH), the integration of which makes it possible to provide automatic validation of interoperability rules to the promoters and participants of the data spaces. It should be noted that the GAIA-X technology stack is under development, and some of the building blocks have been defined but not developed, so the mix with Eclipse perfectly complements the solution.

The core of this architecture is the basis developed in the Horizon Europe DATACELLAR project, one of the “Common Data Spaces of Reference”.

The solution's technological pillars will focus on interoperability and security. EDC facilitates connections between different data platforms by implementing a connector and a federated catalogue. At the same time, Gaia-X ensures that these exchanges comply with European data protection and cybersecurity standards. The proposal adheres to the principles of complete control over the data by the participants by establishing, through policy definitions, who can access their data and under what conditions.

Another main objective of this initiative is security, where advanced mechanisms will be employed to protect data during the exchange, ensure reliable execution, and comply with established data access policies. Open standards will be employed and promoted to facilitate the use of diverse technologies and integrated platforms, thereby achieving

interoperability. The proposed platform will be built using state-of-the-art technologies and architectures that ensure Scalability, Security, and Interoperability.

Finally, a necessary building block under a data space scheme will be developed to ensure the governance framework that supports data sharing: (a) Define the roles and responsibilities of the stakeholders in managing the platform; (b) Establish decision-making processes; (c) Ensure data sovereignty so all data providers retain control over their data and make informed decisions about its use and sharing; (d) Encourage collaboration and communication among stakeholders; and (e) Ensuring compliance with European Data Space principles, including FAIR data management, security, and interoperability.

To access the CRMsDataSpace, a participant needs standardised connectivity methods to ensure interoperability, trust and compliant use of common interaction patterns. Therefore, technical enablement services will be developed to combine identity management, data exchange via a connector and access to compatible digital twins. The steps will be (1) to verify the identity, (2) to connect to the data space, and (3) to exchange standardised data:

- **Signing the Framework Agreements** is essential in providing clarity and legal security for collaboration within the data space. Formalising the data exchange through these agreements lays the foundation for a mutually beneficial relationship. The framework agreements are based on roles ranging from Data Provider/Consumer to Advisory Providers and third-party Application Providers.
- **Registering the company**. An account will be prepared that should be confirmed. A clearing house will validate the data before generating a unique Business Partner Number (BPN). This is central to all further activities within the data space. Once the company receives the BPN, it will have access to the portal. The Gaia-X Digital Clearing House (GXDCH) is a vital component of the Gaia-X initiative, which aims to create a secure and federated data infrastructure in Europe. The GXDCH is a central node for verifying compliance with the Gaia-X rules and enabling organisations to participate in any Gaia-X ecosystems (Automotive, Manufacturing, Aerospace, etc.).
- **Enter the data space**. Setting up the technical environment is needed before exchanging data within the ecosystem. To access CRM dataspace, a participant needs standardised connectivity methods to ensure interoperability, trust, and compliance with common interaction patterns. Therefore, technical enablement services combine identity management, data exchange via a connector and access to compatible digital twins. One thing is crucial: the Eclipse Data Space Connector (EDC). As CRMsDataSpace is open-source, it will be possible to use its service, develop a specific internal solution, or choose one from the data space marketplace.

## 4.3 CHALLENGE 3: To develop multi-sided business solutions for the CRM ecosystem

These multi-side business solutions will leverage network effects, catering to both the "supply" and "demand" of data and creating data-related services to provide or consume data and solve specific ecosystem problems

The CRMS data space ecosystem will be focused on exchanging data on critical raw materials with a universal standard within the EU, enhancing efficiency and leading to significant cost reductions.

Data spaces are a strong and collaborative alternative to platforms. They create “multi-sided business solutions” that rely on network effects, serving both the “supply” and “demand” of data and data-related services. Offering a valuable collection of data resources and services attracts users of data, and a large user base attracts additional data resources and services. However, the challenge in making this work is organisational rather than technical. Creating substantial value requires consensus through collaborative efforts from multiple users.

Beyond the critical objective of secure and sovereign data sharing for value creation, data spaces can be established for different purposes. They can operate as a marketplace, where the space can be seen as a sales channel between data service providers and consumers. Apart from the Core services that enable the essential functions of the CRM ecosystem, such as identity or data discovery on a data space level, the architecture of CRMsDataSpace will be divided into two areas:

- **Enablement Services**: They can be deployed either self-managed by each participant or managed by a service provider to provide or consume data within the CRM data space. To consume data, a GIS web portal will be developed addressing information visualisation (including enhanced visualisation), search, filtering capacity, and download. Complementarily, developing a plugin integrated into the web portal is envisaged to visualise interpreted 3D models of critical raw materials resources. Given the variety and complexity of the data related to the physical environment, the visualisation services will be pretty extensive and can be divided into groups: Spatial context, Layer management; Display of external data (Map tiles: Google Earth, OpenStreetMap, etc.; OGC map services; OpenDat/NetCDF data); Display of tables; Import (csv, shp, wkt, geojson, etc.); Viewing of multimedia files; Measurements, etc. Conversely, a user-friendly interface will be developed to provide data.
- **Business applications**: They solve specific problems within the ecosystem. An app for mobile devices will be developed to optimise, homogenise and guarantee the quality of data collection in the field, automating the acquisition of GPS-georeferenced data (including photographs). The development of this app will

be carried out with the maximum “universality” of implementation in mind, which is why Android is proposed as the Operative System (OS). for its development. Given the limited memory and microprocessor needs, this means a wide range of manufacturers and models of relatively low-cost tablet devices.

Simultaneously, CRMsDataSpace will offer assistance with using the data space or business development consulting for ecosystem participants, such as critical raw materials resource evaluation.

## 4.4 CHALLENGE 4: To implement Artificial Intelligence (AI) reporting tools using Large Language Models (LLMs)

The objective is to automatically generate reports of great interest to data space users, facilitating the search and gathering of information.

A large language model (LLM) is based on transformer networks containing hundreds or billions of parameters. These models have been trained on a gigantic corpus of text, which makes them exceptionally robust in modelling the patterns and structures of human language. Thanks to their vast training datasets, LLMs can learn statistical relationships among words, sentences, and whole texts, allowing them to generate text in a coherent and human-like manner.

Training an LLM is computationally intensive and time-consuming, as it involves tuning millions of parameters based on a vast training corpus. This resource-intensive nature means that training an LLM cannot be done quickly on a personal computer or by outsourcing to cloud services due to prohibitive costs.

LLMs can generalise and apply knowledge to various natural language processing (NLP) tasks thanks to their training on large amounts of text data. However, generalisation also poses challenges. One of the main ones is a bias towards the data used to train them. If an LLM is trained primarily on texts from a specific domain, it may be less effective in other contexts or domains.

In addition, generalisation may be limited by the amount of helpful text available for training, which could lead to diminishing returns as models try to process more data. To address the problem of data bias, these models can be tuned for specific NLP tasks through fine-tuning, a process that involves training the model with a smaller, task-specific corpus. This allows the model to learn domain-specific details and patterns. Specialisation can significantly improve the model's performance on specific tasks, allowing for more accurate and relevant solutions.

Thus, a fine-tuning approach to one or more LLMs will be adopted from the in-house data space and external data sources. This will be particularly useful for improving the

model's generalisation and specialisation in generating user interest reports, allowing it to better adapt to specific tasks and data contexts.

To this end, standard techniques for creating curated data from unstructured data such as PDF files, Word documents or web pages, among others, will be considered, as well as the use of vectorised databases. These organise information into sets of lists, each sorted by a different attribute, allowing related answers to be searched based on their proximity to any number of factors, which is expected to result in greater efficiency, flexibility, accuracy and relevance of the answers generated by the model.

LLMs are, therefore, an ideal tool for generating reports from a template where LLMs are used to populate specific fields, called dynamic fields, of a report template with relevant data, facilitating the automation and creation of reports, reducing the time and effort required compared to manual reporting, and increasing the homogenisation of reports.

Starting from the data space of the project, this line of work within the proposal would include the following stages:

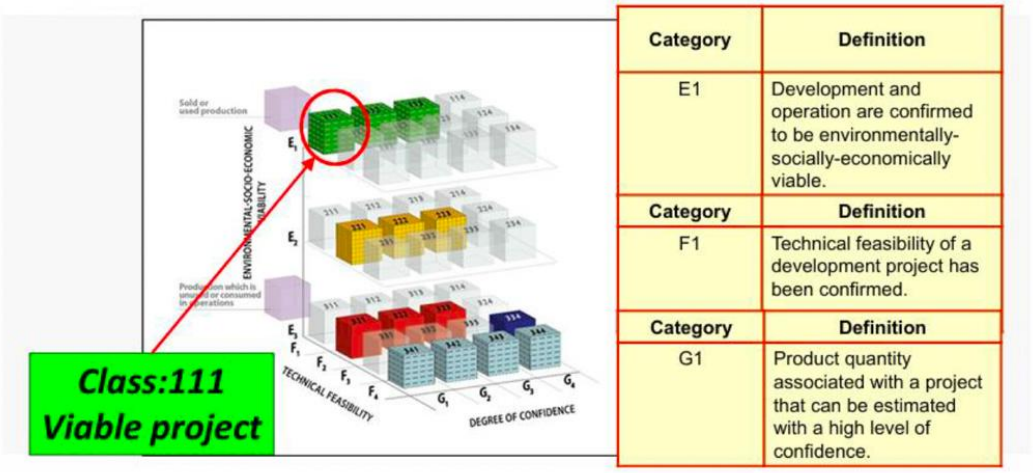
- **Data preparation**: (a) Collecting and organising data relevant to the domain in which the LLM is to specialise. This may involve extracting text from documents, web pages, existing databases, etc.; (b) Separating the data into training, validation, and test sets, helping to evaluate the model's performance objectively; and (c) Using embedding models to convert text into numerical vectors that the LLM can process and semantically capture the text's meaning, which is crucial for fine-tuning.
- **Model Fine Tuning**: (a) Selecting a large language model as the basis for fine-tuning, with many options offered by open-source models, such as those from Hugging Face; (b) Use the prepared data to train the base model so the model learns to adapt to the domain-specific patterns of the training data; and (c) Adjust model parameters, such as learning rate, batch size and the number of epochs, to optimise model performance in fine-tuning. At least 2 LLMs will be trained based on data relevant to the domain.
- **Evaluation and Deployment**: The test data set will be tested to assess how well the model has learned to generalise to new data. This may involve metrics such as accuracy, recall, and F1-score. Once evaluated, the model can be deployed in a production environment and integrated into the system.

At least two language-based models (LLMs) will be adjusted to the data space's contextual information, and at least four report templates will be completed using LLMs.

## 4.5 CHALLENGE 5: To implement an Artificial Intelligence tool to classify CRMs from closed extractive waste facilities according to the United Nations Framework Classification for Resources

The United Nations Framework Classification for Resources (UNFCR) will assist in assessing and monitoring a project's environmental, social, and economic viability, progress, and other criteria.

The UNFCR classification tool developed with AI will assess the availability of mineral projects in a standardised way under defined conditions, with variable conditions in time considered through reconciliation, based on the robust numerical code-based system with supplemental specifications that provide the guidelines for applying the three axes of the rating matrix to mineral projects (Figure 4-3).




| Category | Definition |
|---|---|
| E1 | Development and operation are confirmed to be environmentally-socially-economically viable. |
| **Category** | **Definition** |
| F1 | Technical feasibility of a development project has been confirmed. |
| **Category** | **Definition** |
| G1 | Product quantity associated with a project that can be estimated with a high level of confidence. |

**Figure 4-3. How the UNFCR works[3]**

Classifications will be generated by choosing appropriate combinations of categories and groupings or further subdividing the categories.

Environmental-socioeconomic assumptions will be based on current conditions and realistic assumptions of future conditions. Unless constrained by regulation, assumptions of future market conditions will reflect the view of either the organisation responsible for the evaluation, a competent person or independent evaluator, or an externally published independent view considered a reasonable forecast of future conditions.

[3] European Commission. Commission Staff Working Document on Common European Data Spaces. SWS(2024) 21 final. Brussels, 24.1.2024.

Estimates will only be considered if they are documented in sufficient detail, allowing an independent evaluator or auditor to clearly understand the basis of the estimate and its classification. The proposal will assume this obligation to ensure appropriate internal documentation is generated and kept.

## 4.6 CHALLENGE 6: To deploy an operational Common European Data Space on Critical Raw Materials for the Green Deal supported by CRMsDataSpace

According to the Commission Staff Working Document on Common European Data Spaces4, the European Green Deal recognises data as an essential enabler of the changes needed for a just and green transition of the EU economy and society. The Green Deal data space will support the implementation of European Green Deal policies by providing relevant data (e.g., on air, water, and soil quality related to the zero-pollution strategy) and contribute to a higher level of environmental transparency.

The European strategy for data announced the development of a Common European Green Deal data space to use data's significant potential to support the Green Deal's priority actions on climate change, circular economy, zero pollution, biodiversity, deforestation, and compliance assurance.

GreenData4All is the core initiative for creating the Green Deal data space. In this context, the Commission is reviewing the existing EU rules on environmental geospatial data sharing, leveraging opportunities arising from the legislative instruments that follow the European strategy for data.

## 4.7 Work Plan of the project

The work plan is divided into seven interlinked Work Packages (WPs), which are presented in Figure 4-4:

- WP1: Project management and coordination.
- WP2–WP5: Data acquisition from closed facilities, development of the data catalogue, deploying a minimun viable data space, multi-side business solutions
- WP6: artificial intelligence (AI) tools
- WP7: Communication, dissemination and exploitation.

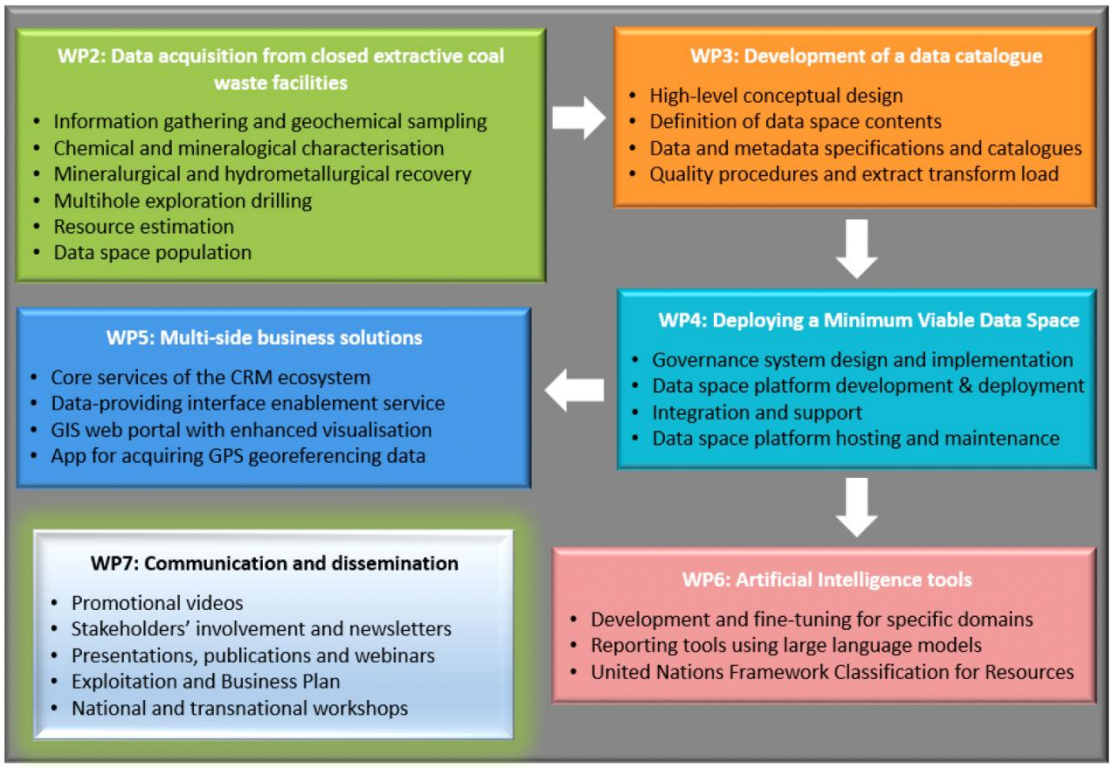



**Figure 4-4. Work plan of the CRMSDataSpace project**

# 5 Impact

## 5.1 Relevant and substantial expected benefits for the European Coal sector

CRMsDataSpace will promote the recovery of critical raw materials from closed extractive coal waste facilities, providing opportunities and attracting investment and jobs to coal regions in transition. This will enable a circular economy within the European Coal Sector, promoting economic growth in areas that would otherwise remain abandoned and would not contribute to maintaining economic activity there.

The proposal will leverage the European coal sector's key competitive advantage, creating sources of wealth from problems such as coal waste heaps that will become solutions for the critical raw materials value chain.

After extracting CRMs, the proposal will contribute to land rehabilitation and ecological restoration in these coal-mining-affected areas, which are often currently degraded. Thus, it aims to accelerate the recovery of degraded and transformed ecosystems to a good ecosystem status, contributing to human well-being.

It will also help protect the most vulnerable workers in the transition by making a vital contribution to the territories expected to be the most negatively impacted by the climate transition. This considers the potential of end-of-life coal mines to extract critical raw materials from closed extractive coal waste facilities, and the need for land rehabilitation and ecological restoration after extraction.

Another pathway towards impact is increasing alternatives in the European Coal sector after the phasing out of coal by repurposing former coal-washing treatment installations into mineralurgical and hydrometallurgical mineral processing facilities focused on critical raw materials.

These repurposed installations shall also concentrate critical raw materials from closed extractive waste facilities other than coal, giving the European Coal sector a business opportunity beyond coal in an area where mining companies have extensive experience and workers with adequate skills for running these mineralurgical and hydrometallurgical mineral processing facilities.

Moreover, specialising in exploring and developing closed extractive waste facilities beyond coal offers another business opportunity for the European Coal sector, given its experience and prior knowledge. The European Union has a legacy of raw materials extraction in many of its regions, and thus, substantial amounts of extractive waste in closed facilities, which, due to their only recent rise in economic importance, have generally yet to be analysed for CRM potential.

The CRMA's measures to establish a common approach to Strategic Projects in the Union that are active in extracting, processing, or recycling strategic raw materials, or in producing relevant substitute materials, support this exciting business opportunity. It is crucial to highlight that, according to the CRMA, to ensure the development of Strategic Projects across the European Union, such projects should benefit from streamlined, predictable permitting procedures and support in accessing finance.

## 5.2 Scale and significance of the project's contribution to the expected outcomes and impacts in the European Coal sector

After successful dissemination and exploitation of project results, CRMsDataSpace expects the following changes in the targeted groups within the European Coal Sector:

1. For more than 1,000 closed extractive coal waste facilities across the European Union, where the particular characteristics of the waste sites or geological conditions make the presence of potentially technically recoverable quantities of critical raw materials likely, analyses are made to determine the approximate quantities and concentrations of all raw materials contained in the extractive waste.
2. Representative geochemical samplings are conducted for more than 450 closed extractive coal waste facilities across the European Union, where available information could indicate the presence of potentially economically recoverable quantities of critical raw materials.
3. More than 250 closed extractive coal waste facilities across the European Union carry out a more detailed sampling with subsequent chemical and mineralogical characterisation involving core logging or equivalent techniques, including classification according to the United Nations Framework Classification for Resources for the most promising critical raw materials.
4. Critical raw materials recovery project developers exploit over 50 closed extractive coal waste facilities within the European Union.
5. More than 15 former coal treatment plants are repurposed into mineral processing facilities focused on critical raw materials.
6. More than 5 mining companies have started businesses related to exploring and developing closed extractive waste facilities other than coal.
7. More than 3,000 workers from the coal mining sector remain employed in the previous activities.

The turnover of the EU coal mining industry was estimated at EUR 27 billion in 2022, according to EURACOAL5. On the other hand, within the EU, around 208,000 people are employed in coal activities, with 76% of these jobs in the mining sector, corresponding to 158,080 people. By scaling up employment figures, the recovery of critical raw materials from closed extractive coal waste facilities may represent 1.9% of the current

coal mining industry turnover, implying an economic activity of more than EUR 0.5 billion.

## 5.3 Requirements and potential barriers

It is well known that the AI and data fields are in a crucial regulatory moment in Europe. As the two of them are key components of the present proposal, the deployed solutions and tasks to be carried out must comply with the present (AI and Data Governance Acts) and future regulations on these matters. As updated regulations are expected due to the evolution of these disciplines, these updates may impose additional requirements on some of the present project's outcomes. 2.2 Measures to maximise impact — Dissemination, exploitation and communication

## 5.4 Target groups, potential users, stakeholders, and multipliers

To maximise the project's impact, the following target groups will be addressed:

Mining industry: Coal mining companies; Operators of coal preparation plants; Senior and Junior general mining companies; Mineral recovery plant operators (mineralurgical and metallurgical); and Mining consultancy companies. Authorities: National Authorities in charge of implementing the CRMA; Authorities from coal regions in transition; National/regional geological surveys and other relevant entities collecting raw materials-related data and information; and EC representatives.

Data services providers: Business solutions, shared industry services, and Data space infrastructure. The scientific community: Mining and geology sciences; Digital science; and Technology and Innovation. Critical raw material suppliers and users.

As multipliers, CRMsDataSpace will address the following ones to reach multiple audiences and stakeholders:

Associations and platforms related to coal mining and mining in general, such as EURACOAL (European Association for Coal and Lignite) and EUROMINES (European Association of Mining Industries, Metal Ores & Industrial Minerals); The Just Transition Platform (JTP); The European Raw Materials Alliance (ERMA); The European Innovation Partnership (EIP) on raw materials; The Raw Materials Information System (RMIS); The United Nations Economic Commission for Europe (UNECE) Group of Experts on Coal Mine Methane and Just Transition; Forums on post-mining perspectives such as the Worl Mining Congress and its International Organizing Committee (WMC-IOC); and EU funded projects that are likely to generate synergies.

# 6 Anticipated Benefits

The project's expected broader scientific effects will stem from increased geological and mining research on CRMs in Europe, leading to the discovery of new deposits. It will also help unleash the enormous potential of data-driven innovation within CRM value chains and allow data from across the EU to be made available and exchanged in a trustworthy and secure manner. Data holders will benefit from a safe and reliable framework for sharing their data to drive innovation. The exploration of implementing the latest AI advancements in raw materials should be expected, including the use of AI for predictive analytics across raw materials value chains, raw materials economic forecasts, and circular economy optimisations, as well as identifying opportunities for AI integration in critical raw materials.

The expected wide-ranging industrial effects of the project concern the need for a robust EU value chain to ensure security of supply and safeguard the functioning of the internal industry. Also, the volatile prices of CRMs create an obstacle for downstream consumers looking to secure stable and predictable prices for key inputs. Moreover, a lack of up-to-date geological information on CRMs in the EU could undermine the development of extraction projects, weakening efforts to decrease supply risk and safeguard the EU industry. Simultaneously, the CRM Common European Data Space will enhance the development of new data-driven products and services for the industry in the EU, increasing its competitiveness in the current global economy.

The expected wider economic effects of the project will arise, in the first place, from the fact that CRMsDataBase will allow potential project promoters to quickly identify facilities with a high potential for economically viable recovery all around Europe without having to correspond with authorities, with uncertainty surrounding access to confidential information. This will facilitate the development of new projects by interested companies and the emergence of junior mining companies that use their capital and efforts to undertake new mining projects, with their financing mainly coming from capital markets, as they are venture capital companies.

Currently, no EU- or Member State-level database reports the volumes of extractive waste deposits (for both closed and active mines), disaggregated by mineral exploitation and/or waste types, and the volumes of the different extractive waste streams. Overall, there are between 150,000 and 500,000 extractive waste deposits in the EU, the majority of which are now inactive, representing accumulations of large amounts of diverse materials, including CRMs and other valuable materials[4].

[4] Blengini, G.A. et al. (2019). Recovery of critical and other raw materials from mining waste and landfills: State of play on existing practices, EUR 29744 EN, Publications Office of the European Union, Luxembourg. https://doi:10.2760/600775, JRC116131.

The previously indicated proportion for the European Coal sector was that 5% of closed extractive coal waste facilities with characteristics or geological conditions that make the presence of potentially technically recoverable quantities of critical raw materials likely would be exploited. Thus, using the lower figure of 150,000 extractive waste deposits and assuming that only 5% of them will have these characteristics, it can be expected that around 7,500 extractive waste facilities would be exploited by critical raw materials recovery project developers. That results in a EUR 75 billion contribution to the EU economy.

Of course, this is accompanied by the subsequent restoration of these areas, with all the societal benefits that this entails. It could be estimated between EUR 30k and 40k of added value per hectare restored[5].

On the other hand, according to the European Commission[6], the value of the data economy exceeded EUR 544 billion in 2023 for the EU27, representing 9.3% growth compared to the previous year. The positive growth trend of the data economy is confirmed by the value of the data market in 2023 for the EU27, which grew at a rate higher than total IT spending, at 11.1% year-on-year, reaching EUR 82 billion. Also, according to the latest estimates, the number of data professionals in the EU27 reached 7.7 million in 2023, accounting for 4.3% of the total workforce, up 4.7% from the previous year.

Assuming that around 30 ecosystems will populate the Common European Data Space in the future, initiating the CRM ecosystem will account for around 3% of the data economy's value, amounting to EUR 16.32 billion in 2023. Extending this value to employment, the number of data professionals in the EU27 may increase by 0.2 million.

[5] Krzemień, A.; Álvarez Fernández, J.J.; Riesgo Fernández, P.; Fidalgo Valverde, G.; Garcia-Cortes, S. Valuation of Ecosystem Services Based on EU Carbon Allowances—Optimal Recovery for a Coal Mining Area. Int. J. Environ. Res. Public Health 2023, 20, 381. https://doi.org/10.3390/ijerph20010381.

[6] European Commission. European Data Market Study 2021-2023. Directorate-General for Communications Networks, Content and Technology. Brussels, 2024.

# 7 Glossary

AI – Artificial Intelligence. Technologies that enable machines to perform tasks that normally require human intelligence.

CRMA – Critical Raw Materials Act. European regulation aimed at securing a sustainable and resilient supply of critical raw materials in the European Union.

CRM – Critical Raw Material. Raw material considered strategically important for the EU economy and the green and digital transitions.

CRMsDataSpace – Project aimed at developing the first Common European Data Space on Critical Raw Materials for the Green Deal.

Data Space – A secure and interoperable framework that enables trusted data sharing among different stakeholders according to common rules and standards.

KPI – Key Performance Indicator. Quantitative measure used to assess progress towards project objectives.

LLM – Large Language Model. Advanced Artificial Intelligence model capable of generating and processing human language.

MVDS – Minimum Viable Data Space. Initial operational version of the data space developed and demonstrated within the project.

RFCS – Research Fund for Coal and Steel, the European funding programme supporting research and innovation projects in these sectors.

UNFC – United Nations Framework Classification for Resources. International system used for the classification and reporting of natural resources and reserves.

WP – Work Package. A major project activity grouping related tasks and deliverables.

Metadata – Structured information describing datasets, facilitating their discovery, management and reuse.

Multi-side Business Solution – A service or business model that creates value for multiple groups of users, such as data providers and data consumers, within the ecosystem.

Stakeholder – Any organisation, authority, company, research institution or individual with an interest in or affected by the project activities and results.

# 8 Annex: State of the art

**1. Relevant ongoing and closed projects**

From the RFCS, the proposal will use the knowledge gained in connection with projects such as:

- **GreenJOBS**, which addresses the recovery of rare earths from coal wastes.
- **WASTE2CRM** assesses the potential and feasibility of using coal mining waste as a source of critical raw materials, specifically rare earth elements.

From Horizon Europe, the proposal will feed from:

- **SCRREEN**, which develops and animates an Expert Network at the EU level covering all the raw materials and their value chains screened in the CRMs assessment.;
- **FutuRaM**, which develops the Secondary Raw Materials knowledge base on the availability and recoverability of secondary raw materials within the EU, with a particular focus on CRMs.
- **GREAT**, which aims to establish the Green Deal Data Space Foundation and its Community of Practice, builds on the European Green Deal and the EU's Strategy for Data.
- **AD4GD's** mission is to co-create and shape the European Green Deal Data Space as an open hub for FAIR data and standards-based services that support the key priorities of pollution, biodiversity and climate change.
- **DS4SSCC** is a pivotal deployment action following the preparatory phase for Data Space for Sustainable and Smart Cities and Communities.
- **FAIRiCUBE**, whose core mission is to enable players from beyond classic Earth Observation (EO) domains to provide, access, process and share gridded data and algorithms in a FAIR and TRUSTable manner.
- **DestinE**, which develops a highly accurate digital model of the Earth (a digital twin of the Earth) to model, monitor and simulate natural phenomena, hazards and related human activities.

**2. Relevant international literature**

1. Blengini, G.A. et al. (2019). *Recovery of critical and other raw materials from mining waste and landfills: State of play on existing practices*, JRC116131, Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/600775,

2. Careddu, N., Dino, G. A., Danielsen, S. W., & Přikryl, R. (2018). *Raw materials associated with extractive industry: An overview*. Resources Policy, 59, 1-6. https://doi.org/10.1016/j.resourpol.2018.09.014

3. Dino, G.A. et al. (2018). *Landfill mining from extractive waste facilities: The importance of a correct site characterisation and evaluation of the potentialities. A case study from Italy*. Resources Policy 59, 50-61. https://doi.org/10.1016/j.resourpol.2018.04.012

4. Kotsev, A., Minghini, M., Tomas, R., Cetl, V., & Lutz, M. (2020). *From spatial data infrastructures to data spaces—A technological perspective on the evolution of European SDIs*. ISPRS International Journal of Geo-Information, 9(3), 176. https://doi.org/10.3390/ijgi9030176

5. Perkins, M. (2023). *Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond*. Journal of University Teaching and Learning Practice, 20(2). https://doi.org/10.53761/1.20.02.07

6. Radu, V. M., Dinca, G., Ivanov, A. A., Szabo, R., & Cetean, V. M. (2024). *New data regarding the identification of critical raw materials recoverable from raw, processed and the waste mining industry materials from Romania*. Environmental Science and Pollution Research, 31(28), 40592-40608. https://doi.org/10.1007/s11356-023-26536-x

7. Rosario-Beltré, A. J., Sánchez-España, J., Rodríguez-Gómez, V., Fernández-Naranjo, F. J., Bellido-Martín, E., Adánez-Sanjuán, P., & Arranz-González, J. C. (2023). *Critical Raw Materials recovery potential from Spanish mine wastes: A national-scale preliminary assessment*. Journal of Cleaner Production, 407, 137163. https://doi.org/10.1016/j.jclepro.2023.137163

8. Tardieu, H. (2022). *Role of Gaia-X in the European data space ecosystem. In Designing Data Spaces: The Ecosystem Approach to Competitive Advantage* (pp. 41-59). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-93975-5

**3. Existing databases**

Many Member States are expected to use existing databases created by their national geological services, in most cases for other purposes, to meet the CRMA's requirements. European-level databases have also been developed under various European projects. The proposal has comprehensively assessed these databases to evaluate their potential.

Consequently, some of the most notable ones are presented here:

**EMODnet Geology**: The European Marine Observation and Data Network provides open access to geological data and metadata held by various geoscience organisations in Europe of the ocean seafloor and its resources, including marine critical minerals. However, in this case, the information available is only the area and the type of potential resource, as shown in Figure 8-1. https://www.emodnet-geology.eu/




Figure 8-1. EMODnet Marine Critical Minerals Map Viewer

**Geoportal of the Czech Geological Survey: It has an inventory of mining waste facilities in the Czech Republic, with a map showing the locations** and details of registered mining waste repositories. The colour differentiation of mining waste repositories distinguishes the categories and levels of exploration. It provides attributes such as coordinates, areas, heights, etc., and photographs. In this case, the information available is just a point on the map and a minimum amount of information, like the coordinates, type of resource, estimated amounts, former operator, etc.

https://mapy.geology.cz/inventarizace_uloznych_mist/

**Geoportal of the Federal Institute for Geosciences and Natural Resources (BGR)**: Data from the fields of soil, geochemistry, geology, geophysics, groundwater, and raw materials, as well as BGR scientific publications, can be searched, viewed, and downloaded. As in the case of EMODnet, the information available is just an area and the type of potential resource, as shown in Figure 8-2.

https://geoportal.bgr.de/mapapps/resources/apps/geoportal/index.html?lang=en#/geoviewer

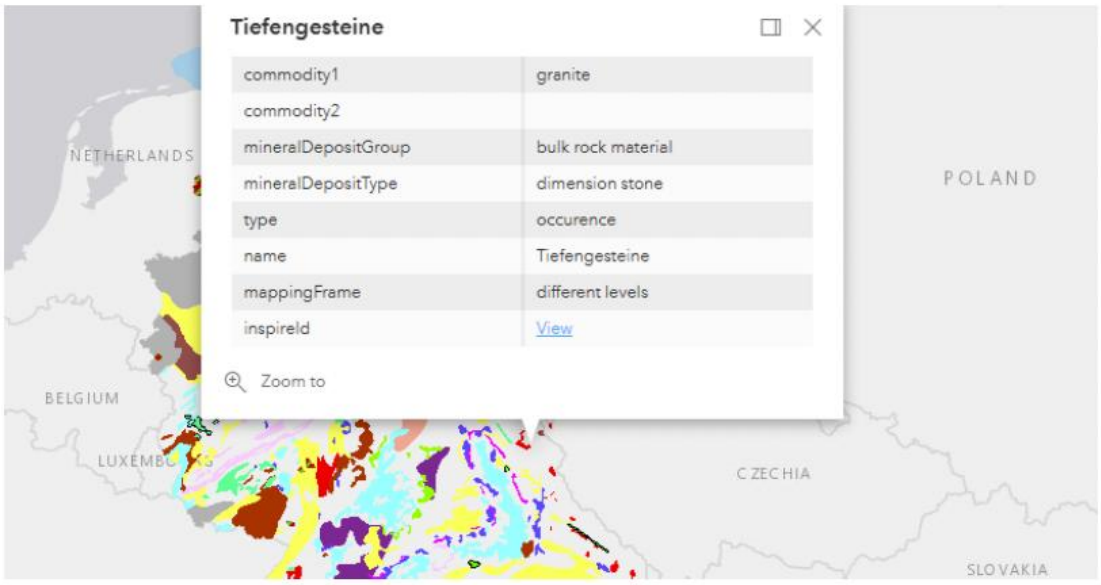



**Figure 8-2. Geoportal of the Federal Institute for Geosciences and Natural Resources**

**Spanish Geological and Mining Institute (IGME)**: This information was obtained by completing the National Inventory of Mining Ponds and Mining Tailings from the IGME from 1983 to 1989, when it stopped. It is a database of sterile deposits, coarse materials poured dry (mine tailings) originating in the mines and quarries, with information related to physical and geological characteristics, type of mining, impact on the environment, the possibility of exploitation, etc. (Figure 8-3).

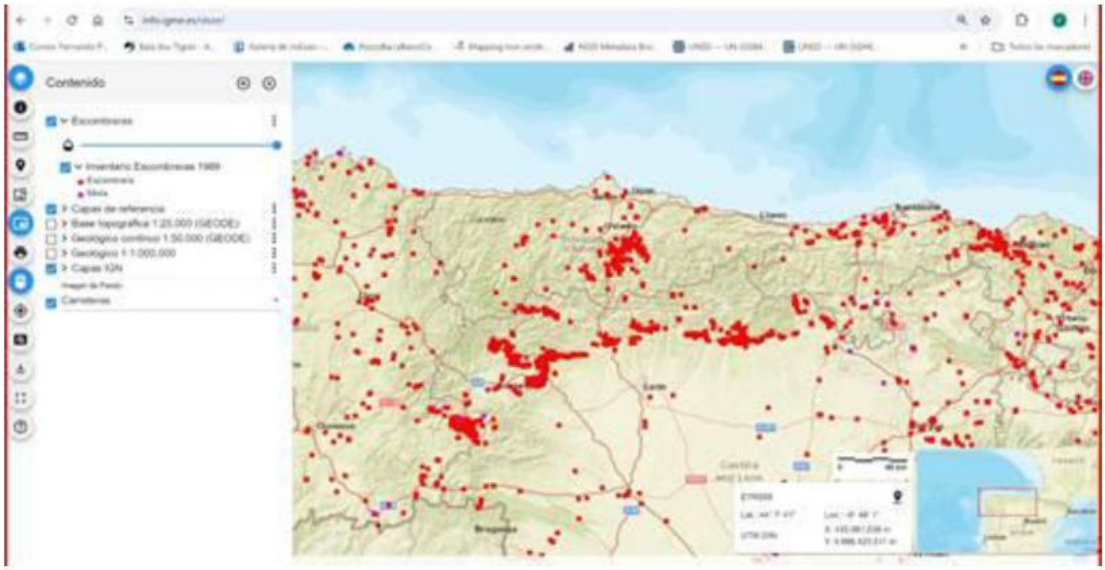



**Figure 8-3. IGME Spanish National Inventory of Mining Tailings**

**Geological Survey Ireland Spatial Resources**: It provides a wide range of datasets, from bedrock, quaternary, groundwater, marine, and geological heritage to geotechnical, minerals, geophysics, geochemistry, and geohazards (Figure 8-4). The Geological Mapping programme creates maps depicting the rocks and subsoils of Ireland's onshore area. It collects new data by field surveying and borehole drilling and combines it with existing mapping, but only to produce map products at various scales.

https://dcenr.maps.arcgis.com/apps/MapSeries/index.html?appid=a30af518e87a4c0ab2fbde2aaac3c228

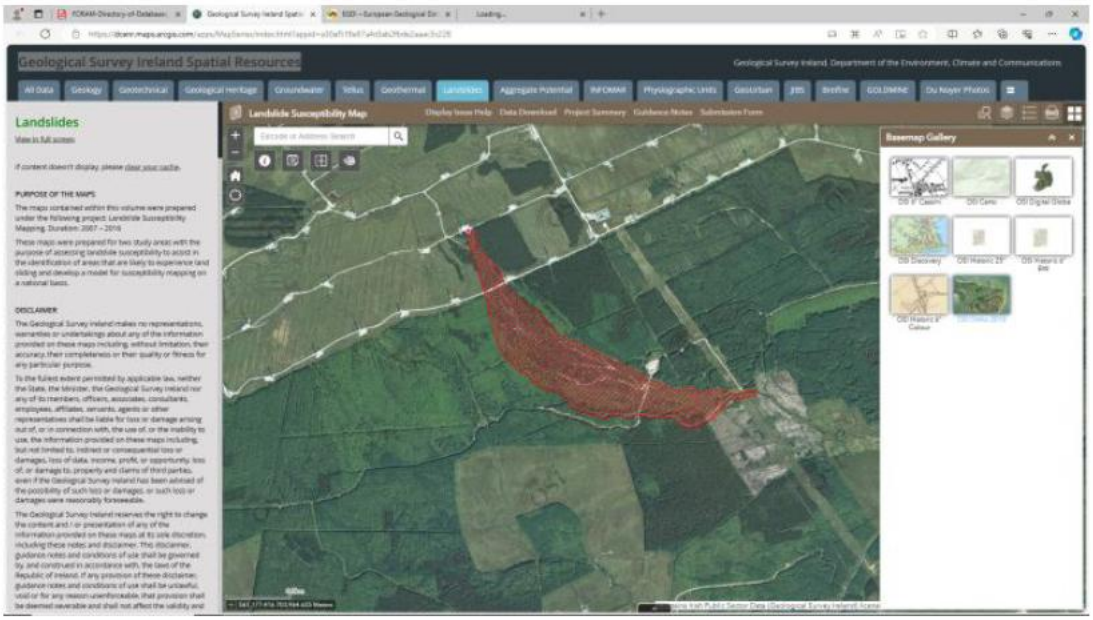



Figure 8-4. Geological Survey Ireland Spatial Resources

**OPI-TPP 2.0 – Poland**: The aim of the project "Extension of the post-mining land management system in the Silesian Province" is to develop and implement a new e-service in the form of a public information system on post-mining areas in the Silesian Province (Figure 8-5). The scope of information and tools to be made available is dedicated to assessing the possibilities and potential for economic and social reuse in these areas. Areas are indicated on the map with specific information, reports, categorisation, development possibilities, photographs, and advanced search capabilities.

https://www.opi-tpp.pl/map

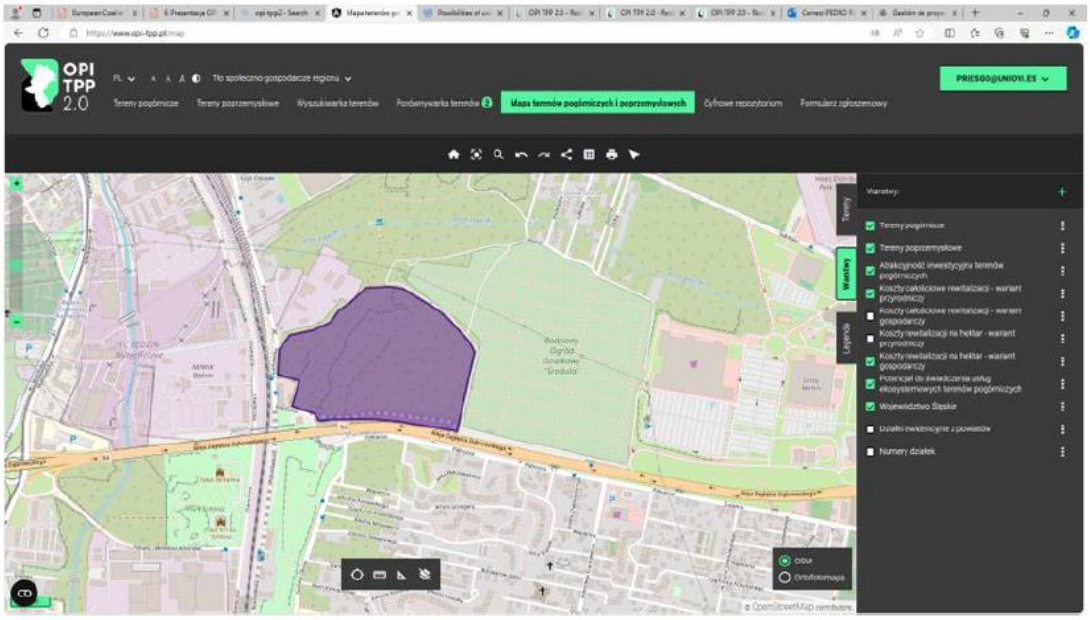



Figure 8-5. OPI-TPP 2.0

**European Geological Data Infrastructure (EGDI)**: It provides access to Pan-European and national geological datasets and services from the Geological Survey Organisations of Europe. One of the datasets presents the critical raw material occurrence points 2023 (EGDI/MIN4EU). However, the occurrences are only represented by dots in the map (which shape represents the critical raw material and the size represents the amount: small, medium, large, very large, occurrence or not specified), and the information is almost nonexistent, e.g., name, critical raw materials, element, mining activity, coordinates and importance (Figure 8-6).

https://data.geus.dk/egdi/?mapname=egdi_new_structure#baslay=baseMapGEUS&extent=2227807.3076396,1742591.2505347852,3182190.367696766,2233759.101526548&layers=egdi_mineraloccurr_critical_raw_materials_2023&filter_0=crm_2023.multi%3D%26commodity_importance.multi%3Dvery%2Blarge%2Bdeposit%26commodity_importance.multi%3Dlarge%2Bdeposit%26mine_status.multi%3D

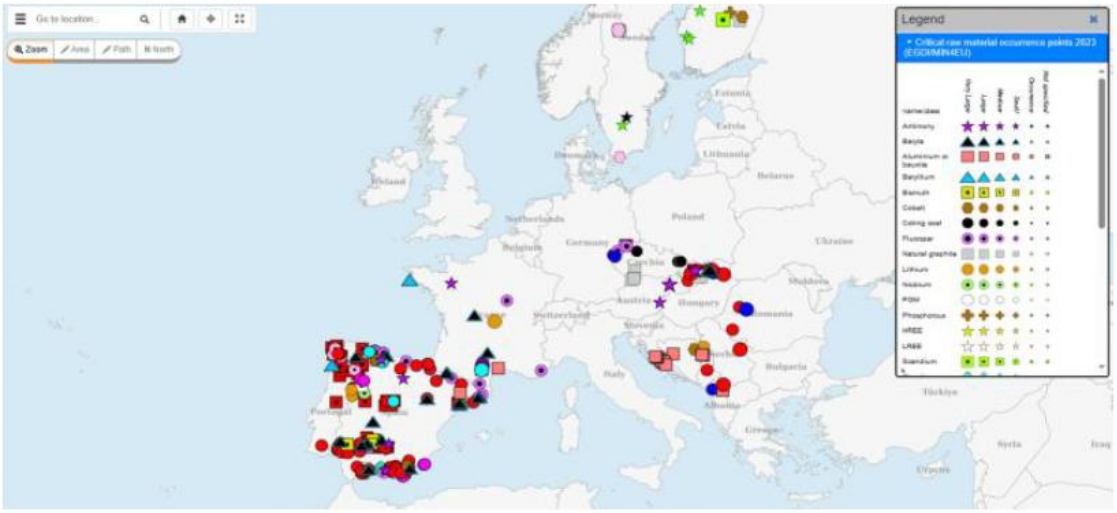



**Figure 8-6. EGDI Critical raw material occurrence points**