Radioactive Waste Management and Decommissioning
NEA/RWM/R(2019)2
March 2022
www.oecd-nea.org

Site Characterisation Library

A Report of the Radioactive Waste Repository Metadata Management (RepMet) Initiative

OECD
BETTER POLICIES FOR BETTER LIVES

NEA
NUCLEAR ENERGY AGENCY

Nuclear Energy Agency

NEA/RWM/R(2019)2

Unclassified **English text only**

1 March 2022

NUCLEAR ENERGY AGENCY

Cancels & replaces the same document of 22 February 2022

Radioactive Waste Management Committee

Site Characterisation Library

A Report of the Radioactive Waste Repository Metadata Management (RepMet) Initiative

This document is available in PDF format only.

JT03490465

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 38 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

NUCLEAR ENERGY AGENCY

The OECD Nuclear Energy Agency (NEA) was established on 1 February 1958. Current NEA membership consists of 34 countries: Argentina, Australia, Austria, Belgium, Bulgaria, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, Norway, Poland, Portugal, Romania, Russia, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Commission and the International Atomic Energy Agency also take part in the work of the Agency.

The mission of the NEA is:

– to assist its member countries in maintaining and further developing, through international co-operation, the scientific, technological and legal bases required for a safe, environmentally sound and economical use of nuclear energy for peaceful purposes;

– to provide authoritative assessments and to forge common understandings on key issues as input to government decisions on nuclear energy policy and to broader OECD analyses in areas such as energy and the sustainable development of low-carbon economies.

Specific areas of competence of the NEA include the safety and regulation of nuclear activities, radioactive waste management and decommissioning, radiological protection, nuclear science, economic and technical analyses of the nuclear fuel cycle, nuclear law and liability, and public information. The NEA Data Bank provides nuclear data and computer program services for participating countries.

Acknowledgements

The Nuclear Energy Agency (NEA) wishes to express its gratitude for the preparation of the report to all the members of the Integration Group for the Safety Case (IGSC) Radioactive Waste Repository Metadata Management (RepMet) initiative. Special thanks to:

- Wilfried Albert (National Cooperative for the Disposal of Radioactive Waste [Nagra], Switzerland)
- Massimo Ciambrella (NEA)

And:

- László Sőrés (BTIX, Hungary – Consultant)
- Simon Lambert (United Kingdom – Consultant)
- Mark Thorley (United Kingdom – Consultant)

Table of contents

Tables

Figures

Boxes

Executive summary

The Radioactive Waste Repository Metadata Management (RepMet) initiative was launched in 2014 by the Integration Group for the Safety Case (IGSC) of the OECD Nuclear Energy Agency (NEA) Radioactive Waste Management Committee (RWMC). RepMet analysed and investigated the application of metadata, a fundamental tool of modern data and information management, within national programmes for radioactive waste repositories. This analysis concluded that there is a great need and potential for metadata management and harmonisation.

A special characteristic of radioactive waste repositories is the long time between construction and closure of the facility – typically periods in excess of one hundred years. This means that systems handling data and relevant supporting information (metadata) will, in all likelihood, go through technological and other changes: data media and the data themselves may become unreadable and programs handling such data may become obsolete. In addition, successive generations of workers will perform tasks on the site during this period with a high probability that not all knowledge will be handed down through the generations. Therefore, the data handling operations must enable the long-term, intergenerational reliability and usability of data.

Given this challenge, the main aim of RepMet has been to formulate a consistent set of guiding principles for capturing and generating metadata, in order to enable national programmes to create sets of metadata that can be used to manage their repository data, information and records in a way that is both harmonised internationally and suitable for long-term management and utilisation in safety cases and elsewhere.

RepMet has produced five interrelated reports that discuss the key aspects of data and related metadata for selected scientific and technical topics involved in the life cycle of a radioactive waste repository. These reports include, and are underpinned by, three technical libraries containing high-level conceptual data models (CDMs), descriptions of data entities, attributes, associated metadata and controlled dictionaries. The libraries can be used independently of each other; however, utilising all of the libraries and the approach outlined in these documents helps provide the additional benefit of a uniform approach to metadata management.

This document, the "Site Characterisation Library", is the second of these five reports. It supports an associated technical library dealing with data and related metadata that are considered during the characterisation of a site investigated and surveyed for suitability for radioactive waste disposal purposes, leading up to site selection.

The Site Characterisation Library has two principal aims:

- to show how the use of appropriate metadata can support the long-term management of the "core information" that is acquired during the characterisation of a site investigated and surveyed for suitability for radioactive waste disposal purposes, leading up to site selection;

- to provide application examples about how implementing the metadata-based techniques can support the long-term management of the "core information".

Several Radioactive Waste Management Organisations (RWMOs) and research laboratories from NEA countries were involved in the RepMet initiative: Andra (France), Enresa (Spain), JAEA (Japan), Nagra (Switzerland), RWM/NDA (United Kingdom), NWMO (Canada), ONDRAF/NIRAS (Belgium), Posiva (Finland), PURAM (Hungary), Sandia National Laboratories (United States), SKB (Sweden) and SÚRAO (Czech Republic).

It is hoped that RepMet activities will contribute to easing the data management burden on individual RWMOs and will be a move towards interoperability and harmonisation. A joint set of principles, controlled dictionaries, data model libraries, etc., can facilitate data exchange with common stakeholders such as international peer review groups, NGOs and regulators. This approach should allow less mature programmes to benefit from the advances made by other sister organisations. Adoption of RepMet's CDMs can contribute to improving the quality and cost-effectiveness of an RWMO's data and metadata management activities.

RepMet does not intend to promote any commercial products or services for managing data or information.

List of abbreviations and acronyms

Andra	Agence nationale pour la gestion des déchets radioactifs (National Radioactive Waste Management Agency, France)
CDM	Conceptual data model
CGI	Commission for the Management and Application of Geoscience Information
Enresa	Empresa Nacional de Residuos Radioactivos S.A. (National Radioactive Waste Company, Spain)
ERD	Entity Relationship Diagram
GeoSciML	GeoScience Markup Language
GML	Geography Markup Language
IFEP	International Features, Events and Processes (NEA)
IGSC	Integration Group for the Safety Case (NEA)
INSPIRE	Infrastructure for Spatial Information in Europe
ISO	International Organization for Standardization
ISO/TC211	ISO Technical Committee 211 on Geographic information/Geomatics
JAEA	Japan Atomic Energy Agency
MRMS	Minnesota Recordkeeping Metadata Standard
Nagra	National Cooperative for the Disposal of Radioactive Waste (Switzerland)
NEA	Nuclear Energy Agency
NUMO	Nuclear Waste Management Organization of Japan
NWMO	Nuclear Waste Management Organization (Canada)
O&M	Observations and Measurements Standard
OECD	Organisation for Economic Co-operation and Development
OGC	Open Geospatial Consortium
ONDRAF/NIRAS	National Agency for Radioactive Waste and Enriched Fissile Material (Belgium)
Posiva	Expert organisation in nuclear waste management (Finland)
PURAM	Public Limited Company for Radioactive Waste Management (Hungary)
RDF	Resource Description Framework
RepMet	Radioactive Waste Repository Metadata Management (NEA)

RMDC	Recordkeeping Metadata Development Committee (MRMS)
RWM	Radioactive Waste Management
RWMC	Radioactive Waste Management Committee (NEA)
RWM/NDA	Radioactive Waste Management/Nuclear Decommissioning Authority (United Kingdom)
RWMO	Radioactive Waste Management Organisation
SensorML	Sensor Model Language
SKB	Nuclear Fuel and Waste Management Company (Sweden)
SKOS	Simple Knowledge Organization System
SÚRAO	Radioactive Waste Repository Authority (Czech Republic)
SWE	Sensor Web Enablement
URL	Universal Resource Locator
W3C	World Wide Web Consortium
XML	Extensible Markup Language

1. Introduction

1.1. The aim of the RepMet initiative

In order to support their operational, pre- or post-closure safety cases and other requirements, Radioactive Waste Management Organisations (RWMOs) manage very large amounts of data that they both produce and receive. A special characteristic of radioactive waste repositories is the long time between construction and closure of the facility – typically periods in excess of one hundred years. This means that systems handling data and relevant supporting information (metadata) will, in all likelihood, go through technological and other changes: data media and the data themselves may become unreadable and programs handling such data may become obsolete. In addition, successive generations of workers will perform tasks on the site during this period with a high probability that not all knowledge will be handed down through the generations. Therefore, the data handling operations of RWMOs must enable the long-term, intergenerational reliability and usability of data.

Given this challenge, the main aim of RepMet has been to formulate a consistent set of guiding principles for capturing and generating metadata, in order to enable national programmes to create sets of metadata that can be used to manage their repository data, information and records in a way that is both harmonised internationally and suitable for long-term management and utilisation in safety cases and elsewhere.

Box 1.1: What is RepMet?

The Radioactive Waste Repository Metadata Management (RepMet) initiative was launched in 2014 by the Integration Group for the Safety Case (IGSC) of the Radioactive Waste Management Committee (RWMC) at the Nuclear Energy Agency (NEA). RepMet analysed and investigated the application of metadata, a fundamental tool of modern data and information management, within national programmes for radioactive waste repositories. Based on this analysis it was realised that there is a great need and potential for metadata management and harmonisation.

Several RWMOs and research laboratories from NEA countries were involved in the RepMet initiative: Andra (France), Enresa (Spain), JAEA (Japan), Nagra (Switzerland), NUMO (Japan), ONDRAF/NIRAS (Belgium), Posiva (Finland), PURAM (Hungary), RWM/NDA (UK), Sandia National Laboratories (United States), SKB (Sweden) and SÚRAO (Czech Republic).

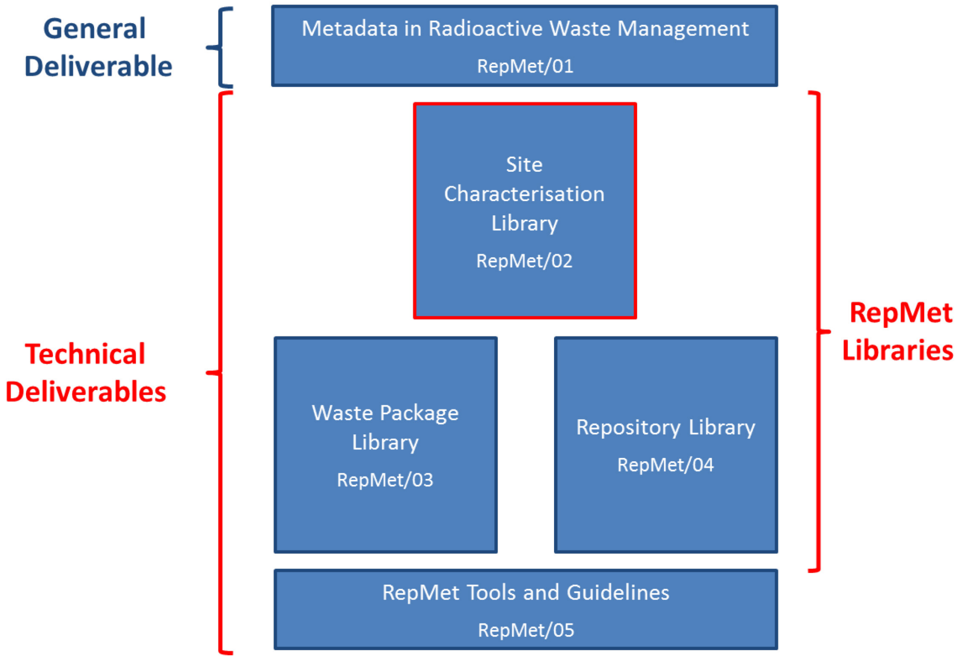
RepMet does not intend to promote any commercial products or services for managing metadata.

1.2. The products of the RepMet initiative and their intended audience

RepMet has produced five key interrelated documents, as summarised in Figure 1.1.

The information provided within these documents is primarily aimed at RWMOs that are considering developing information systems or establishing knowledge management practices related to geological disposal, or that are planning to renew or update their existing data management practices. This information is intended to be sufficiently generic to enable it to be adapted by almost any RWMO. The information may also be of use for other disciplines such as those related to developing inventory and decommissioning models.

Figure 1.1: The RepMet document family



Source: NEA, 2019.

The five documents[1] are as follows:

RepMet/01 – *Metadata for Radioactive Waste Management* (NEA, 2018) provides an overview of metadata and its application within RWMOs, discusses issues around the implementation of metadata, and outlines the outputs of RepMet and how they may be used. It also provides specific recommendations concerning metadata for RWMOs.

The three reports identified as "RepMet Libraries" are more technically detailed. They discuss the key aspects of data and related metadata for selected scientific and technical topics involved in the life cycle of a radioactive waste repository. The reports include, and are underpinned by, three technical libraries, containing high-level conceptual data models, descriptions of data entities, attributes, associated metadata and other relevant information, and are ready to support the activities of RWMOs. The libraries can be used independently of each other; however, utilising all of the libraries and the approach outlined in these documents helps provide the additional benefit of a uniform approach to metadata management.

1. The documents are available in electronic format on the RepMet webpage on the NEA website. See www.oecd-nea.org/jcms/pl_61001.

RepMet/02 – "Site Characterisation Library" (this document) deals with data and related metadata that are considered during the characterisation of a site investigated and surveyed for suitability for radioactive waste disposal purposes, leading up to site selection.

RepMet/03 – "Waste Package Library" (NEA, 2021a) deals with data and related metadata about packaged waste and spent nuclear fuel that, after proper treatment and conditioning processes, are ready for final disposal at the repository.

RepMet/04 – "Repository Library" (NEA, 2021b) deals with data and related metadata linked to the engineered structures and waste acceptance requirements of radioactive waste repositories.

RepMet/05 – "RepMet Tools and Guidelines" (NEA, 2021c) supports the libraries, providing a number of tools, methods, guidelines and approaches that were either used in developing the libraries or will be useful for RWMOs when adopting and implementing the libraries.

Table 1.1: Intended audiences for RepMet documents

Deliverable	Primary audience	Secondary audience
RepMet/01 – *Metadata for Radioactive Waste Management*	RWMO Managers and Decision Makers: • What metadata are and why they are valuable to their organisations; • Issues to consider in metadata implementation, and how RepMet proposals may be adopted; • High-level recommendations on metadata adoption and implementation at an organisational level. Information Systems Developers: • Awareness of benefits and risks in metadata implementation projects; • Identification of possible designated communities for metadata use.	Local and international regulators Other concerned authorities: • Awareness of role of metadata in ensuring audit trails and long-term reliability of data, information and records. Non-specialist audiences: • Understanding of best practices in information handling in RWM, and expectations on what information should be available over the long term.

Table 1.1: Intended audiences for RepMet documents (Continued)

Deliverable	Primary audience	Secondary audience
RepMet/02 – Site Characterisation Library RepMet/03 – Waste Package Library RepMet/04 – Repository Library	Information Systems Developers: • Reusable data models and controlled dictionaries developed and validated by RepMet. RWMO Engineers: • Awareness of attributes of interest to information systems for long-term access and use; • Agreed vocabulary for international harmonisation of terms.	Academics: • Current best practice in metadata modelling for RWMOs, as basis for further development in future.
RepMet/05 – RepMet Tools and Guidelines	Information Systems Developers: • Tools and techniques for use during the implementation process; • Recommended existing standards and how they may be applied.	RWMO managers or decision makers interested in technical aspects (e.g. data modelling).

Source: NEA, 2019.

The documents are primarily designed for use by personnel in RWMOs, regardless of whether they have a strong background or not in such areas as database management, database development, data modelling or any other area of information and/or computing systems. The documents provide high-level overviews and summaries suitable for RWMO managers and decision makers, and include more detailed, implementation specific information targeted at information system developers working within a RWMO environment. See Table 1.1 for details of the intended audiences.

1.3. An introduction to RepMet/02 – Site Characterisation Library

The Site Characterisation Library presents a collection of data and metadata models to describe the "geological and geophysical characterisation of the repository site". The library includes examples of the application of the data model to real-world geological and environmental properties that RWMOs need to investigate during the characterisation of the site of a disposal facility.

Site characterisation requires knowledge about rock stability and isolation properties that are controlled by surface and underground conditions close to the repository, and in the wider environment. A complex data model of the geological environment including information on composition, structure and processes is needed. This knowledge is developed by using a wide range of geoscience disciplines, such as geology, geophysics, hydrogeology, geochemistry and analysis of satellite imagery. There is also a requirement for ongoing monitoring, repeating observations on a regular basis. This involves handling large amounts of data with a varied structure and content that can arise for methodological, historical and/or practical reasons.

In the last decade, considerable effort has been devoted to the harmonisation of data related to knowledge about the natural and artificial environment. A major European activity is INSPIRE (Infrastructure for Spatial Information in Europe) (European Union, 2007) that integrates 34 environmental themes into one uniform conceptual data model (CDM). It is based on open standards and harmonised web services.

The high-level CDM proposed for the Site Characterisation Library is largely based on INSPIRE principles and data models and covers the geo-scientific aspects of site characterisation relevant for nuclear waste disposal facilities. The controlled dictionaries, relating to geology, geophysics and environmental monitoring, are also based on INSPIRE.

A common structure is used for the Site Characterisation Library, the Waste Package Library (NEA, 2021a) and the Repository Library (NEA, 2021b). For the Site Characterisation Library this is as follows:

- Chapter 2 introduces the metadata standards that RepMet investigated and used for the development of the Site Characterisation Library. For this library, it is mainly based on the data specification that the Infrastructure for Spatial Information in Europe (INSPIRE) provides.
- Chapter 3 reports the CDMs that RepMet developed for the Geology, Geophysics and Environmental Monitoring domains, based on the INSPIRE data models. For each domain, the chapter includes an overview of the proposed entities with definitions taken from the original INSPIRE documentation (in some cases simplified or extended as required), an Entity Relationship Diagram (ERD) and a detailed list of proposed attributes with definitions.
- Chapter 4 presents real-world application examples of the CDMs introduced in Chapter 3.
- Chapter 5 reports the controlled dictionaries that the RepMet group identified for the geological and geophysical characterisation of the repository site and for environmental monitoring, defining harmonised terminology in these three areas.
- Chapter 6 closes the report and provides considerations for future work.

These chapters contain information about metadata-based standards and techniques, including the Observation and Measurement (O&M) standard and the Minnesota Recordkeeping Metadata Standard (MRMS). These are at an introductive level only, and for more details see the RepMet Tools and Guidelines report (NEA, 2021c).

2. Review of existing standards

2.1. Scope

Prior to the establishment of the RepMet initiative there were a lack of national and international metadata standards that specifically supported the management of radioactive waste. This lack of domain specific standards led the IGSC to establish the RepMet initiative within the NEA framework with the remit to investigate the use of metadata to support and improve the management of data and information related to radioactive waste management.

The Site Characterisation Library is a technical report designed to show the application of metadata tools and techniques within the area of geological and geophysical characterisation of the repository site for radioactive waste management. Rather than developing new standards, the RepMet group reviewed a range of existing national and international geoscience data standards and selected a number that can be usefully applied in the area of radioactive waste management.

2.2. Metadata standards in the Site Characterisation Library

There are a variety of data standards related to geoscience, and it is difficult to encompass all of them in a single project like RepMet. In the last decade, considerable effort has been devoted to the harmonisation of data related to knowledge about the natural and artificial environment. A major European activity is INSPIRE (Infrastructure for Spatial Information in Europe) (European Union, 2007) that addresses 34 spatial data themes needed for environmental applications. It is based on open standards and harmonised web services, and makes use of four main pillars called Foundation Schemas. Foundation Schemas encompass standards that are grouped by topic or the standardisation body. These are as follows:

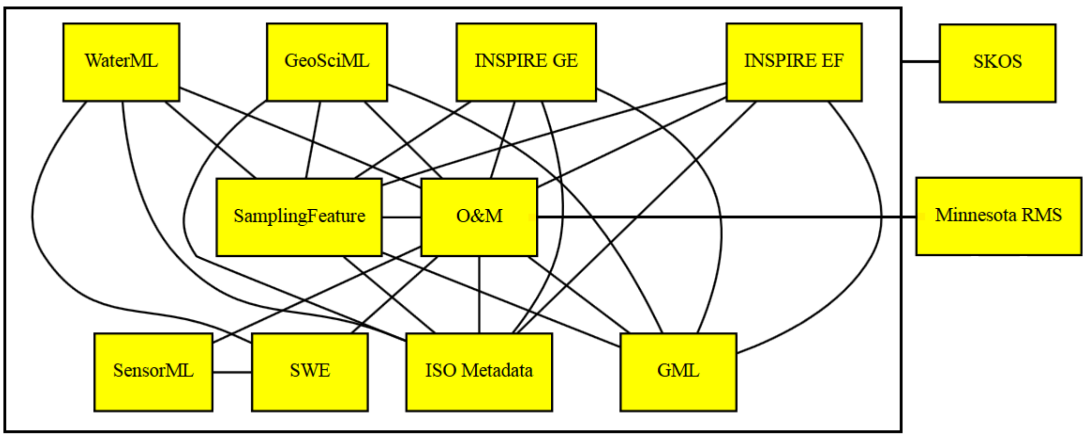
- ISO/TC211 series;
- Open Geospatial Consortium (OGC) standards;
- GeoSciML; and
- EarthResourceML.

The ISO/TC211 series is a group of 38 fundamental standards elaborated by the Technical Committee 211 (TC211) on “Geographic information/Geomatics” of the International Organization for Standardization (ISO). They cover various fields of applications ranging from the basic, like language and country codes, spatial and temporal schema, to the complex, like web services. OGC standards contain essential schemas defining basic data types for numerical data storage and time series. GeoSciML is the product of a pre-INSPIRE standardisation project carried out by the Commission for the Management and Application of Geoscience Information (CGI), an international organisation of the geology community. EarthResourceML is a small package partly based on GeoSciML that is used by the Minerals sector.

The four INSPIRE Foundation Schemas group 40 individual standards and other schemas that are interconnected, referencing each other, and form a large system with a complex hierarchy. For site characterisation purposes a well-defined subset of these 40 standards can be selected and used in data modelling.

From all the standards contributing to the 4 INSPIRE pillars, 11 have been selected. Based on the importance of official record-keeping in repository management the Minnesota Recordkeeping Metadata Standard (MRMS) has been also used. This subset of standards and their relations are shown on Figure 2.1. The main aim of this figure is to show the complexity of the connections between the selected standards, and the central role of O&M and the O&M Sampling Feature schema. SKOS is also fundamental, as it is extensively used by all the INSPIRE standards and many other applications to represent relationships between terms in controlled dictionaries. To keep readability connections from SKOS to the other standards are omitted on the Figure 2.1.

Figure 2.1: Standards to be considered in site characterisation data modelling



Source: NEA, 2019.

The following paragraphs give an overview of these individual standards that are most relevant in site characterisation.

2.2.1. ISO Metadata

ISO19115 (Geographic information – Metadata) (Open Geospatial Consortium, n.d.) is a generic spatial metadata standard. In INSPIRE and OGC compliant web services, the XML implementation of the standard ISO19139 is used.

2.2.2. GeoScience Markup Language (GeoSciML)

GeoScience Markup Language (GeoSciML) (CGI, n.d.) is a model of geological features commonly described and portrayed in geological maps, cross sections, geological reports and databases. It describes a logical model and GML/XML encoding rules for the exchange of geological map data, geological time scales, boreholes and metadata for laboratory analyses. It includes a lightweight model, used for simple map-based applications; a basic model, aligned on INSPIRE, for basic data exchange; and an extended model to address more complex scenarios.

2.2.3. Observations and Measurements (O&M) standard

Observations and Measurements (O&M) was developed by the OGC and is implemented as the ISO standard 19156 "Geographic information – Observations and Measurements" (Cox [ed.], Open Geospatial Consortium Inc., 2013). The O&M standard defines a conceptual data model to represent and encode observations, and, as an extension, measurements based on sampling. It structures and arranges the data and metadata in an organised and regular way that helps to maintain and preserve the information associated with an observation. Though originally developed for geographic information, this standard is generic and can be applied to many types of observational data, including those related to radioactive waste management.

The O&M standard is based on the concept of an "observation". This is any act of observing a property of a feature of interest resulting in the estimation of a value, and involving application of specified processes such as measurement and numerical simulation.

A key element of the O&M standard is that instead of using different data models for different kind of observations, a single conceptual data model works for all. This enables improved interoperability between different information systems, and makes database development easier. The diversity of real-world observations, including those from the management of radioactive waste, is implemented through the adoption of appropriate and specific controlled dictionaries[2] for the elements of the O&M conceptual model. The addition of new fields or new types of observation is undertaken by updating these controlled dictionaries. RepMet has adopted specific controlled dictionaries supporting the geophysical properties of sites being investigated and surveyed for their suitability for radioactive waste disposal purposes. These are reported in Chapter 5.

O&M SamplingFeature

The basic O&M standard can be used to model any kind of direct observations. In order to use the O&M standard to model indirect observations, it is necessary to adopt the "Sampling Feature" extension (Open Geospatial Consortium, 2007a). Indirect observations include, for example, observations involving sampling techniques where a measurement can be used to infer the value of a property of a feature of interest. These sampling features provide a link between features of technical interest and the observation metadata. Sampling features are often related to each other, as parts of associated sets or complexes, through sub-sampling etc.

2.2.4. Sensor Model Language (SensorML)

Sensor Model Language (SensorML) (Open Geospatial Consortium) was created by the sensor community and provides standard models and an XML encoding for describing

2. A controlled dictionary (also called a controlled vocabulary) is a collection of agreed terms that a community or an organisation uses, manages and maintains in a controlled way within a particular domain of interest. The terms will refer to entities within the domain and their attributes. All terms in a controlled dictionary have unambiguous and non-redundant identification, and may be connected to each other through clearly defined relationships declaring, for example, that one term is broader than another. There may also be multilingual labels for terms, allowing consistent usage in different languages.

sensors and measurement processes. SensorML can be used to describe a wide range of sensors, including both dynamic and stationary platforms and both in situ and remote sensors. This makes SensorML a relevant tool for describing observation procedures when required by the O&M standard.

2.2.5. Sensor Web Enablement (SWE)

Sensor Web Enablement (SWE) is part of INSPIRE OGC Foundation Schemas. The Sensor Web Enablement Suite has been designed for the encoding and provision of observational data. It provides suitable data types for measures, counts, quantities, categories, and other useful elements to enable online encoding and data exchange.

2.2.6. Geography Markup Language (GML)

Geography Markup Language (GML), ISO19136, (Portele [ed.], Open Geospatial Consortium, 2007b) is an XML encoding for expressing geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the internet.

2.2.7. WaterML

WaterML 2.0 is a standard information model for the representation of water observation data, intended to allow the exchange of such data sets across information systems.

2.2.8. INSPIRE Geology data theme

In the INSPIRE context the Geology data theme (INSPIRE Thematic Working Group Geology, 2013) can be seen as a “reference data theme” as it provides information for several other INSPIRE data themes. It is composed of the following sub-themes:

- **Geology:** provides basic knowledge about the physical properties and composition of geologic materials (rocks and sediments), their structure and their age as depicted in geological maps, as well as landforms (geomorphological features). The model also covers boreholes − another important source of information for interpreting the subsurface geology.
- **Hydrogeology:** describes the flow, occurrence and behaviour of water in the subsurface environment. The two basic elements are the rock system (including aquifers) and the groundwater system (including groundwater bodies). Man-made or natural hydrogeological objects/features (such as groundwater wells and natural springs) are also included.
- **Geophysics:** focuses on the availability and location of key geophysical features. It includes metadata on high rank gravity, magnetic and seismological stations that are part of international and national observation networks as well as metadata on 2D and 3D seismic measurements that are most often requested by third party users. It also provides collective metadata on gravity, magnetic and airborne geophysical campaigns that cover large areas and provide basic geological information for scientific research and more detailed applied studies, e.g. exploring earth resources (hydrocarbons, mineral deposits, ground water, geothermal energy).

Most INSPIRE themes have two data models: core and extension. Core schemas contain the most important elements of the themes that are obligatory for data providers. Implementing Rules of the INSPIRE directives are based on the core model. Extension schemas are more sophisticated, providing optional technical elements to help experts to share details in a harmonised way. For example, the use of Observations and Measurements is typical in the extension schemas.

2.2.9. INSPIRE Environmental Monitoring Facilities spatial data theme

The INSPIRE Environmental Monitoring Facilities spatial data theme (INSPIRE Thematic Working Group Environmental Monitoring Facilities, 2013) includes the environmental monitoring facility as a spatial object and the data obtained through observations and measurements taken at this facility, encoded using the O&M standard. This information is complemented by further administrative information pertaining to the facility and activities undertaken there such as networks or programmes. The Environmental Monitoring Facilities theme is cross-cutting to environmental domains; thus, the generic model allows the necessary freedom to bring in thematic-specific needs while keeping a shared data structure.

2.2.10. Minnesota Recordkeeping Metadata Standard (MRMS)

Minnesota Recordkeeping Metadata Standard (MRMS) is a standard that the Recordkeeping Metadata Development Committee of the US State of Minnesota developed to facilitate record management at the governmental level, releasing version 1.3 of MRMS in 2015. It shares many of its elements with other metadata standards, such as the Dublin Core[3] and ISO 19115[4]. Apart from information on format, location and access, MRMS provides elements to describe responsible parties, management, preservation history, and all administrative details that are relevant for the life cycle of material in hardcopy, analogue or digital form (see RMDC, 2015, for more details).

RepMet considered that the use of MRMS for record-keeping at the government level provides a good basis for record-keeping within RWMOs. It has also been tested and used by PURAM (Hungary). RepMet therefore adopted and adapted the MRMS to provide the framework for record-keeping integrated into the metadata models that the initiative developed. The integration of the MRMS and the O&M metadata models provides a global schema to encode observations and their records.

2.2.11. Dublin Core

The Dublin Core Metadata Element Set is a simple model for generic purpose metadata. There is significant overlap with ISO19115. For geographic Information in INSPIRE the ISO19115 Metadata standard and its XML implementation ISO19139 is used.

2.2.12. Simple Knowledge Organization System (SKOS)

Simple Knowledge Organization System (SKOS) is a World Wide Web Consortium (W3C) standard to represent "knowledge organisation systems" – taxonomies, thesauri and other types of structured controlled dictionaries.

3. Dublin Core: The Dublin Core Metadata Initiative provides a simple model for general-purpose metadata. There is significant overlap with ISO19115. See DCMI Usage Board, http://dublincore.org.

4. ISO 19115 (Geographic information – Metadata) is a generic spatial-metadata standard (Open Geospatial Consortium [n.d.], retrieved from www.iso.org/obp/ui/#iso:std:iso:19115:-1:ed-1:v1:en).

SKOS is built on Resource Description Framework (RDF), a W3C standard for the conceptual description or modelling of information about web resources – that is, anything that can be identified through a location on the Web. SKOS is a RDF vocabulary to create RDF databases about structured controlled dictionaries with their hierarchical and semantic relations.

3. Conceptual data models (CDMs)

3.1. Scope

A "data model" is an abstract representation of the structure and logical organisation of a database. A database is an organised collection of data about a specific business area of interest, the geological and geophysical characterisation of the repository site in the case of the Site Characterisation Library.

A conceptual data model (CDM) is a high-level data model intended to represent the semantics of an entire domain of interest. It describes the organisation and the structure of a database in terms of objects of interest (i.e. entities) together with their descriptive characteristics (i.e. attributes) and logical associations among them (i.e. relationships). A CDM is not related to the software and hardware used to create a database, so allowing database designers to represent data independently from information systems. For more details, the "RepMet Tools and Guidelines" report (NEA, 2021c) contains a specific section dedicated to data modelling.

For the development of the Site Characterisation Library, RepMet used geoscience domain CDMs derived from the standards introduced in Chapter 2. The implementation of the CDMs presented in this section make it possible to provide datasets that are fully compliant with these standards. The entities are simplified from those described in the standards, however they still have all the attributes that are sufficient for generating valid XML.

3.1.1. Level of data modelling in the library

Because it has a strict connection with INSPIRE, the data modelling for the Site Characterisation Library is closer to a logical level rather than a conceptual model. However, RepMet provided a conceptual level vision of the data models adopted from INSPIRE. There is a key difference between the Site Characterisation Library and the other two libraries. The entities in the Site Characterisation Library have attributes based on the existing INSPIRE controlled dictionaries, whereas the entities of the Waste Package and Repository Libraries have attributes directly structured in controlled dictionaries that the RepMet group originally developed.

Box 3.1: Complex and simple attributes in a Site Characterisation Library

Many of the attributes in these conceptual models are "complex", and further data modelling is required to fully describe them at the level of individual "simple" attributes. For example: "GenericName", "Identifier" and "Entity_Name" are all complex attributes with embedded structures; however, to describe them in detail is outside the scope of this document. The referenced standard schemas provide directions for the next phase of design. For illustrative purposes the three example complex attributes can be further detailed as follows:

GenericName is used to refer to dictionary items. Is consists of a namespace and an identifier code.

Identifier is used for global identification with namespace, local-id and version.

"Entity_Name" is a link to the substructure of the referred entity.

Attribute and property are used as synonyms in the text. Simple attribute and simple property both refer to name-value pairs with simple numerical, logical or textual value. Complex attributes and complex properties refer to entities that have their own simple attributes (or properties) but not explained in detail on the current level of modelling.

Box 3.2: RepMet Terminology - Attribute vs Data

In the terminology adopted by RepMet, "attribute" and "data" are two sides of the same coin. "Attribute" is a property or a characteristic of interest in a database, "data" is the value (for example, a number, a function, a string or some text) that an attribute can assume. For example, if "hydraulic conductivity" is the attribute about a certain rock, then "8x10-7 m/s" may be the numeric data value.

3.2. Site Characterisation Library CDMs

This section presents the CDMs for the domains reported in Figure 3.1: geology (including hydrogeology and geomorphology), geophysics and environmental monitoring. These CDMs are based on the INSPIRE data specifications, and entity definitions and attribute descriptions are also taken from the INSPIRE documentation.

3.2.1. Geology CDM

The Geology CDM is based on the INSPIRE Geology core schema. For some entities, GeoSciML 3.2 was also considered[5]. However, the highly detailed modelling from the GeoSciML model was ignored, and only key attributes are included.

5. For example, "Contact" – a feature to describe geological interfaces – is missing from the INSPIRE core model, but is an important feature in 3D geological modelling. The link between Geology and O&M, realised by the "samplingFrame" attribute of the "Mapped Feature" entity is also taken from GeoSciML 3.2.

Table 3.1: Geology CDM – Entity definitions

Entity	Definition
Geologic Feature	Abstract class that holds common properties of Geologic Units and Geologic Structures (such as name and geologic history).
Geologic Unit	A volume of rock that has distinct characteristics. It can be a lithologic, crono-, biostratigraphic unit, etc. The composition of a Geologic Unit can be described by proportion and lithology value (e.g. 60% sand, 40% clay) in INSPIRE. GeosciML has a more sophisticated submodel for describing composition called EarthMaterial.
Composition Part	It describes the composition of the geologic unit.
Geologic Structure	Abstract class that holds common properties of shear displacement structures, folds and contacts.
Shear displacement structure	Defined as a brittle to ductile style structure along which displacement has occurred.
Fold	A fold is defined as one or more systematically curved layers, surfaces or lines in a rock body.
Contact	Interface between two distinct geologic units.
Geologic Event	An identifiable event during which one or more geological processes act to modify geological entities. Geological age is modelled using GeologicEvent.
Mapped Feature	A spatial representation of a Geologic Feature. It provides a link between a notional feature (description package) and one spatial representation of it, or part of it (exposures, surface traces and intercepts, etc.) which forms the specific bounded occurrence, such as an outcrop or map polygon[6].

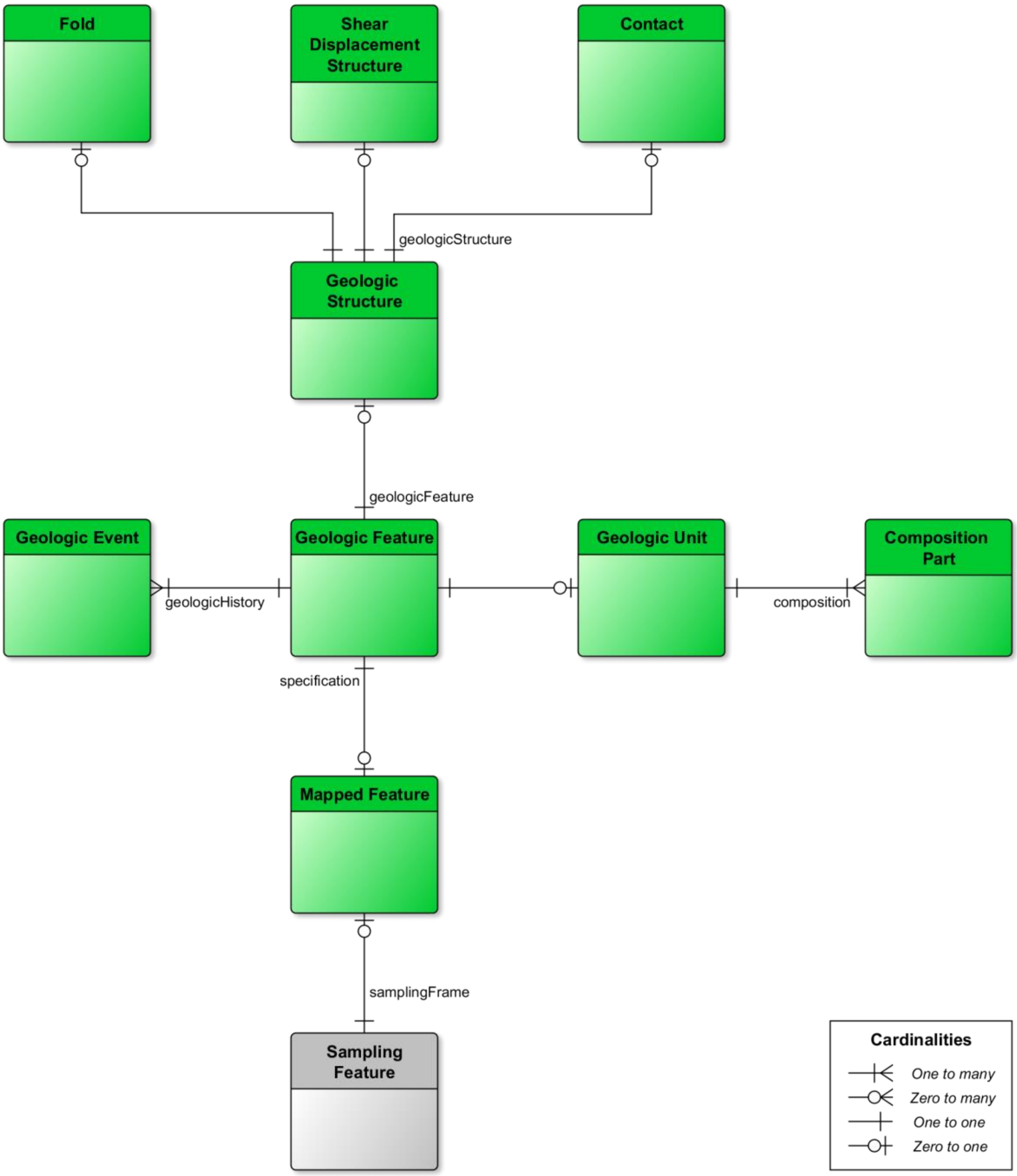
Source: NEA, 2019.

Table 3.1 reports the entity definitions, and Figure 3.1 illustrates the Geology CDM in the Entity Relationship Diagram (ERD)[7] format. The attributes of these entities are described in Tables 3.2 to 3.10. For each attribute, the tables specify the cardinality (i.e. the number of possible occurrences corresponding to each entity instance), the type of data associated with the attribute and a brief attribute description. Colour coding is used to help in the visualisation of the connection between the tables.

6. “Mapped Features” are specified as Geologic Features of different types, like Geologic Unit, Contact, or Geologic Structure. Structures can be Folds or Faults (Shear Displacement Structure). Mapped Features may be associated with Sampling Features that hold observation metadata and results.

7. More details on Entity Relationship Diagrams can be found in the “RepMet Tools and Guidelines” (NEA, 2021c).

Figure 3.1: Geology CDM – Entity Relationship Diagram



Source: NEA, 2019.

Box 3.3: What are cardinalities?

Each relationship in an ERD has an associated cardinality. This describes the minimum and the maximum numbers of occurrences of one entity that may be related to a single occurrence of the other entity. Because all relationships are bidirectional, cardinality must be defined in both directions for every relationship. The cardinality is represented on the ERD through the use of a graphical marker on each end of the relationship as is shown in the legend in Figure 3.1. Cardinalities are explained in more detail in Chapter 2 of the RepMet Tools and Guidelines report (NEA, 2021c).

Table 3.2: Attributes for "Geologic Feature" entity

Entity	Attribute	Cardinality	Data Type	Description
Geologic Feature	**identifier**	[1..1]	*Identifier*	Unique identifier for the geologic feature.
	name	[1..1]	*GenericName*	Human readable name of the geologic feature.
	geologicHistory	[1..N]	*"Geologic Event"*	Reference to the attributes of the "Geologic Event" entity (see Table 3.3.).

Source: NEA, 2019.

Table 3.3: Attributes for "Geologic Event" entity

Entity	Attribute	Cardinality	Data Type	Description
Geologic Event	**name**	[1..1]	*GenericName*	Human readable name for the geologic event.
	eventEnvironment	[1..1]	*GenericName*	Geologic environment of the event.
	eventProcess	[1..N]	*GenericName*	Type of geologic process.
	olderNamedAge	[1..1]	*GenericName*	Start time of process.
	youngerNamedAge	[1..1]	*GenericName*	End time of process.

Source: NEA, 2019.

Table 3.4: Attributes for "Geologic Unit" entity

Entity	Attribute	Cardinality	Data Type	Description
Geologic Unit	***geologicFeature***	[1..1]	*"Geologic Feature"*	Reference to the attributes of the "Geologic Feature" entity (see Table 3.2).
	geologicUnitType	[1..1]	*GenericName*	Type of geologic unit.
	composition	[1..N]	*"Composition Part"*	Reference to the attributes of the "Composition Part" entity (see Table 3.5).

Source: NEA, 2019.

Table 3.5: Attribute for "Composition Part" entity

Entity	Attribute	Cardinality	Data Type	Description
Composition Part	material	[1..1]	*GenericName*	The material that comprises part or all of the geologic unit.
	role	[1..1]	*GenericName*	The relationship of the composition part to the geologic unit composition as a whole.
	proportion	[1..1]	*QuantityRange*	Quantity that specifies the fraction of the geologic unit composed of the material.

Source: NEA, 2019.

Table 3.6: Attributes for "Geologic Structure" entity[8]

Entity	Attribute	Cardinality	Data Type	Description
Geologic Structure	***geologicFeature***	[1..1]	*"Geologic Feature"*	Reference to the attributes of the "Geologic Feature" entity (see Table 3.2).
	identifier	[1..1]	*Identifier*	Unique identifier for the geologic structure.

Source: NEA, 2019.

Table 3.7: Attributes for "Shear Displacement Structure" entity

Entity	Attribute	Cardinality	Data Type	Description
Shear Displacement Structure	***geologicStructure***	[1..1]	*"Geologic Structure"*	Reference to the attributes of the "Geologic Structure" parent entity (see Table 3.6).
	faultType	[1..1]	*GenericName*	Type of shear displacement structure.

Source: NEA, 2019.

Table 3.8: Attributes for "Fold" entity

Entity	Attribute	Cardinality	Data Type	Description
Fold	***geologicStructure***	[1..1]	*"Geologic Structure"*	Reference to the attributes of the "Geologic Structure" parent entity (see Table 3.6.)
	profileType	[1..1]	*GenericName*	Type of fold profile.

Source: NEA, 2019.

8. "Geologic Structure" is an abstract entity. It serves as a link between a specific type of geological structure (Fold, Shear Displacement or Contact) and a specific Geologic Feature. For this reason, it does not contain specific attributes.

Table 3.9: Attributes for "Contact" entity

Entity	Attribute	Cardinality	Data Type	Description
Contact	***geologicStructure***	[1..1]	*"Geologic Structure"*	Reference to the attributes of the "Geologic Structure" parent entity (see Table 3.6).
	contactType	[1..1]	*GenericName*	Type of contact.
	contactCharacter	[1..N]	*GenericName*	Character of the boundary, as opposed to its type.
	Orientation	[1..N]	*CGI_PlanarOrientation*	Orientation of the contact surface.

Source: NEA, 2019.

Table 3.10: Attributes for "Mapped Feature" entity

Entity	Attribute	Cardinality	Data Type	Description
Mapped Feature	***specification***	[1..1]	*"Geologic Feature"*	Reference to the attributes of the "Geologic Feature" entity (see Table 3.2).
	samplingFrame	[0..1]	*"Sampling Feature"*[9]	Reference to the "SamplingFeature" entity (see section 4.4.8.1 of "RepMet Tools and Guidelines" [NEA, 2021c]).

Source: NEA, 2019.

3.2.2. Geophysics CDM

The Geophysics CDM is based on the INSPIRE Geology/Geophysics core and extension schemas. Table 3.11 reports the entity definitions, and Figure 3.2 illustrates the Geophysics CDM in ERD format. The attributes of these entities are described in Tables 3.12 to 3.23. For each attribute, the tables specify the cardinality (i.e. the number of possible occurrences corresponding to each entity instance), the type of data associated with the attribute and a brief attribute description.

9. This data type is illustrated in Table 4.9 of the "RepMet Tools and Guidelines" (NEA, 2021c).

Table 3.11: Geophysics CDM – Entity definitions

Entity	Definition
Geophysical Object	Generic class for geophysical objects. It models single geophysical entities that are used for spatial sampling either by means of data acquisition or data processing.
Geophysical Measurement	Generic spatial object type for geophysical measurements. Geophysical measurements collect data outside or on the boundary of the observed spatial domain.
Geophysical Station	Geophysical measurement spatially referenced to a single point location.
Geophysical Profile	Geophysical measurement spatially referenced to a curve.
Geophysical Swath	Geophysical measurement spatially referenced to a surface.
Geophysical Model	Geophysical object that is created as a result of geophysical data processing or interpretation.
Curve Model	Geophysical model that represents a curve coverage of some geophysical properties.
Surface Model	Geophysical model that represents a surface coverage of some geophysical properties.
Solid Model	Geophysical model that represents a solid coverage of some geophysical properties.
Geophysical Object Set	Generic class for collections of geophysical objects. It is a set of geophysical objects that are grouped by some common property.
Campaign	Geophysical activity extending over a limited time range and limited area for producing similar geophysical measurements, processing results or models.
Project	Geophysical activity extending over a longer time range and larger area, containing any number of campaigns or subprojects.

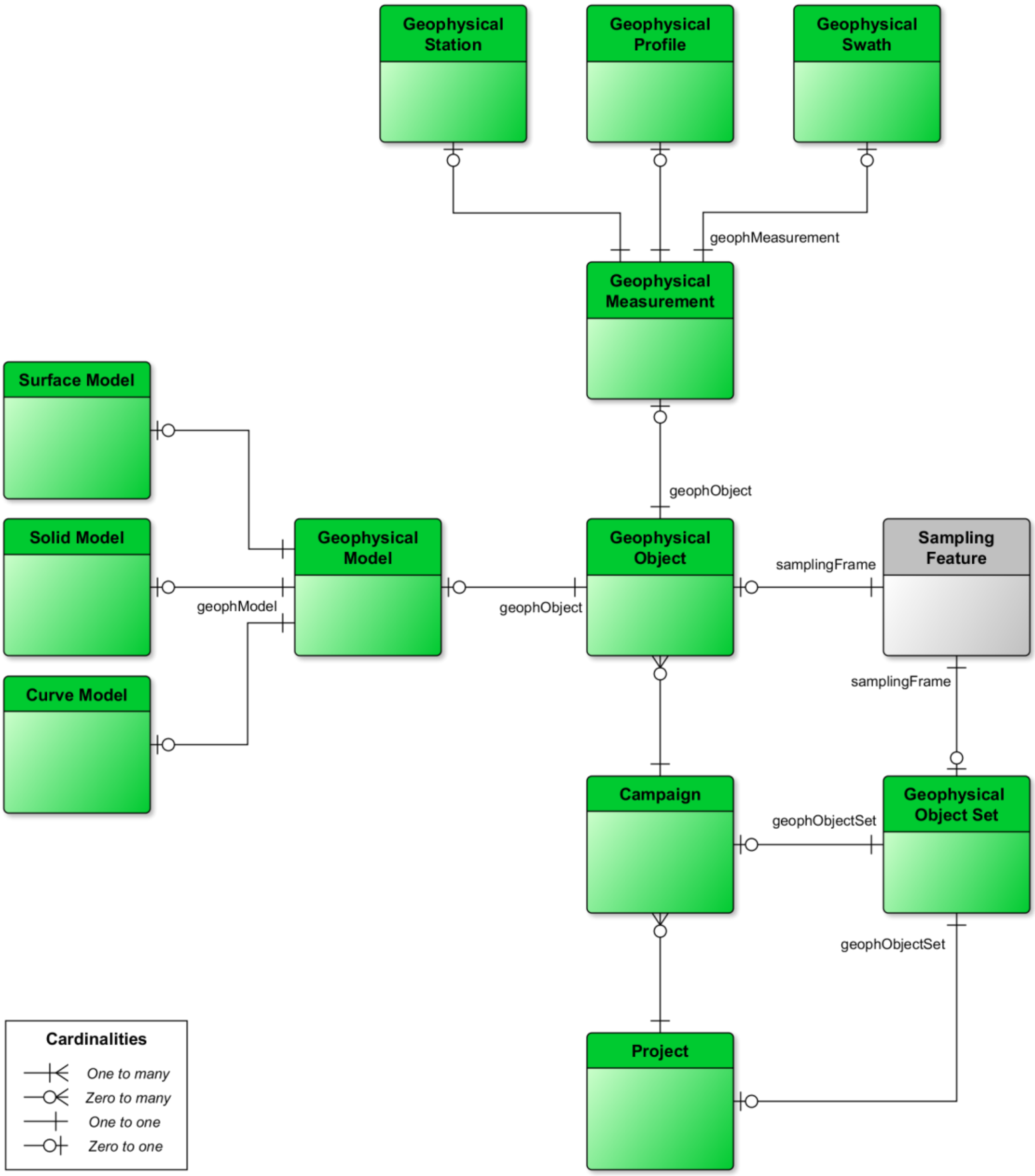
Source: NEA, 2019.

The geophysical features, which are shown in Figure 3.2 are directly linked to the O&M Sampling Feature extension (see Section 2.2.3). Geophysics Objects may be grouped to form Geophysics Object Sets like Campaigns and Projects, and can be used to model the hierarchy of site exploration activity. The two main Geophysics Object types are Geophysics Measurement and Geophysics Model. Both are classified by their sampling geometry. Subtypes of measurements and models cover the complete range of geophysical features that can be found in practice. Colour coding is used to help in the visualisation of the connection between the tables.

The distinction between measurement and model is not always straightforward:

- Measurement data are usually used by domain experts as input for further processing or interpretation. Examples include seismic field data and borehole logs with raw geophysical profiles.
- Geophysical models represent spatial distribution of physical or geophysical properties within the observed spatial domain. Examples include seismic depth sections and composite logs.

Figure 3.2: Geophysics CDM – Entity Relationship Diagram



Source: NEA, 2019.

Table 3.12: Attributes for "Geophysical Object" entity

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Object	***samplingFrame***	[1..1]	*"Sampling Feature"*	Reference to the attributes of the "Sampling Feature" entity (see section 4.4.8.1 of the "RepMet Tools and Guidelines" [NEA, 2021c]).
	projectedGeometry	[1..1]	*Geometry*	2D projection of the feature to the ground surface (as a representative point, curve or bounding polygon) to be used by an INSPIRE view service to display the spatial object location on a map.
	largerWork	[1..1]	*Identifier*	Identifier of a larger work dataset, typically a campaign or project.
	verticalExtent	[1..1]	*EX_VerticalExtent*	Physical extent or estimated exploration depth.

Source: NEA, 2019.

Table 3.13: Attributes for "Geophysical Measurement" entity

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Measurement	***geophObject***	[1..1]	*"Geophysical Object"*	Reference to the attribute of the "Geophysical Object" parent entity (see Table 3.12).
	platformType	[1..1]	*GenericName*	Platform from which the measurement was carried out.
	relatedNetwork	[1..1]	*GenericName*	Name of a national or international observation network to which the facility belongs, or to which measured data are reported.

Source: NEA, 2019.

Table 3.14: Attributes for "Geophysical Station" entity

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Station	***geophMeasurement***	[1..1]	*"Geophysical Measurement"*	Reference to the attributes of the "Geophysical Measurement" parent entity (see Table 3.12).
	stationType	[1..1]	*GenericName*	Type of geophysical station.
	stationRank	[1..1]	*GenericName*	Geophysical stations may be part of a hierarchical system. Rank is proportional to the importance of a station.

Source: NEA, 2019.

Table 3.15: Attributes for "Geophysical Profile" entity

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Profile	***geophMeasurement***	[1..1]	*"Geophysical Measurement"*	Reference to the attributes of the "Geophysical Measurement" parent entity (see Table 3.13).
	profileType	[1..1]	GenericName	Type of geophysical profile.

Source: NEA, 2019.

Table 3.16: Attributes for "Geophysical Swath" entity

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Swath	***geophMeasurement***	[1..1]	*"Geophysical Measurement"*	Reference to the attributes of the "Geophysical Measurement" parent entity (see Table 3.13).
	swathType	[1..1]	GenericName	Type of areal measurement.

Source: NEA, 2019.

Table 3.17: Attributes for "Geophysical Model" entity[10]

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Model	***geophObject***	[1..1]	*"Geophysical Object"*	Reference to the attributes of the "Geophysical Object" parent entity (see Table 3.12).
	identifier	[1..1]	*Identifier*	Unique identifier for the geophysical model.

Source: NEA, 2019.

Table 3.18: Attributes for "Curve Model" entity

Entity	Attribute	Cardinality	Data Type	Description
Curve Model	***geophModel***	[1..1]	*"Geophysical Model"*	Reference to the attributes of the "Geophysical Model" parent entity (see Table 3.17).
	modelType	[1..1]	*GenericName*	Type of curve model.

Source: NEA, 2019.

10. "Geophysical Model" is an abstract entity. It serves as a link between a specific type of geophysical model (Surface, Solid or Curve Model) and a specific Geophysical Object. For this reason it does not contain specific attributes.

Table 3.19: Attributes for "Surface Model" entity

Entity	Attribute	Cardinality	Data Type	Description
Surface Model	**geophModel**	[1..1]	*"Geophysical Model"*	Reference to the attributes of the "Geophysical Model" parent entity (see Table 3.17).
	modelType	[1..1]	*GenericName*	Type of surface model.

Source: NEA, 2019.

Table 3.20: Attributes for "Solid Model" entity

Entity	Attribute	Cardinality	Data Type	Description
Solid Model	**geophModel**	[1..1]	*"Geophysical Model"*	Reference to the attributes of the "Geophysical Model" parent entity (see Table 3.17).
	modelType	[1..1]	*GenericName*	Type of solid model.

Source: NEA, 2019.

Table 3.21: Attributes for "Geophysical Object Set" entity

Entity	Attribute	Cardinality	Data Type	Description
Geophysical Object Set	**samplingFrame**	[1..1]	*"Sampling Feature"*	Reference to the attributes of the "Sampling Feature" entity of the O&M CDM (see section 4.4.8.1 of the "RepMet Tools and Guidelines" [NEA, 2021c]).
	projectedGeometry	[1..1]	*Geometry*	2D projection of the feature to the ground surface (as a representative polygon) to be used by an INSPIRE view service to display the spatial object location on a map.
	largerWork	[1..1]	*Identifier*	Identifier of a larger work dataset, typically a project.
	verticalExtent	[1..1]	*EX_VerticalExtent*	Physical extent or estimated exploration depth.

Source: NEA, 2019.

Table 3.22: Attributes for "Campaign" entity

Entity	Attribute	Cardinality	Data Type	Description
Campaign	***geophObjectSet***	[1..1]	*"Geophysical Object Set"*	Reference to the attributes of the "Geophysical Object Set" parent entity (see Table 3.21).
	campaignType	[1..1]	*GenericName*	Type of geophysical campaign.
	surveyType	[1..1]	*GenericName*	Type of geophysical method used in the campaign.
	client	[1..1]	*"Responsible Party"*[11]	Key responsible party for which data were created. Reference to the attributes of the "Responsible Party" entity of the O&M CDM (see section 4.4.7 *of* "RepMet Tools and Guidelines" [NEA, 2021c]).
	contractor	[1..1]	*"Responsible Party"*	Key responsible party by which data were created. Reference to the attributes of the "Responsible Party" entity of the O&M CDM (see section 4.4.7 *of* "RepMet Tools and Guidelines" [NEA, 2021c]).

Source: NEA, 2019.

11. This data type is illustrated in Table 4.7 of "RepMet Tools and Guidelines" (NEA, 2021c).

Table 3.23: Attributes for "Project" entity

Entity	Attribute	Cardinality	Data Type	Description
Project	***geophObjectSet***	[1..1]	*GeophObjectSet*	Reference to the attributes of the "Geophysical Object Set" parent entity (see Table 3.21).
	principalInvestigator	[1..1]	*ResponsibleParty*	Key responsible party for conducting research. Reference to the attributes of the "Responsible Party" entity of the O&M CDM (see section 4.4.7 *of* "RepMet Tools and Guidelines" [NEA, 2021c]).

Source: NEA, 2019.

Box 3.4: How to handle other types of measurement

Alongside geophysical measurements, other types of measurements also play an important role in site characterisation. The general hydrogeological conditions of the host rock are often examined by hydraulic tests which are an essential element in site characterisation. Hydraulic tests are carried out in boreholes and are often preceded or followed by traditional borehole logging.

Using the Sampling Feature CDM, a hydraulic test can be described as a special kind of observation that may be linked to both a geological (borehole) and a geophysical (borehole logging) entity. Hydraulic tests are usually carried out in multiple boreholes simultaneously, and this can be described by the use of Sampling Feature Complexes. The function of different boreholes can be defined by the role attribute using controlled dictionary items like "pumping well" or "observation well" etc. Process parameters are useful in documenting measurement details.

As well as hydraulic tests other "non-standard" measurements can also be described in a standard way by using the generic O&M Sampling Feature model together with controlled dictionaries of domain specific terms.

3.2.3. Environmental Monitoring CDM

The Environmental Monitoring CDM is based on the INSPIRE "Environmental monitoring facilities" theme that can be utilised for any generic monitoring application including the monitoring of radioactive waste repositories.

This INSPIRE theme has a complex data model, designed in part to support legal and reporting obligations which are less important in the context of site characterisation. Therefore, only part of the full data model of the INSPIRE theme was used to develop the CDM for Environmental Monitoring.

The entities that RepMet selected for the Environmental Monitoring CDM are described in Table 3.24 and Figure 3.3 illustrates the Environmental Monitoring CDM in the ERD format. The attributes are explained from Table 3.25 to Table 3.30. For each attribute, the tables specify the cardinality (i.e. the number of possible occurrences corresponding to each entity instance), the type of data associated with the attribute and a brief attribute description. Colour coding is used to help in the visualisation of the connection between the tables.

In the original INSPIRE model, the "Monitoring Feature" entity is linked directly to the "Observation" entity of the O&M standard. To have a consistent conceptual model in RepMet, a slightly different arrangement was used. Because in the Environmental Monitoring domain observation always means spatial sampling, the use of the "Sampling Feature" model is highly recommended. For this reason, the "Monitoring Feature" entity is connected to the Sampling Feature.

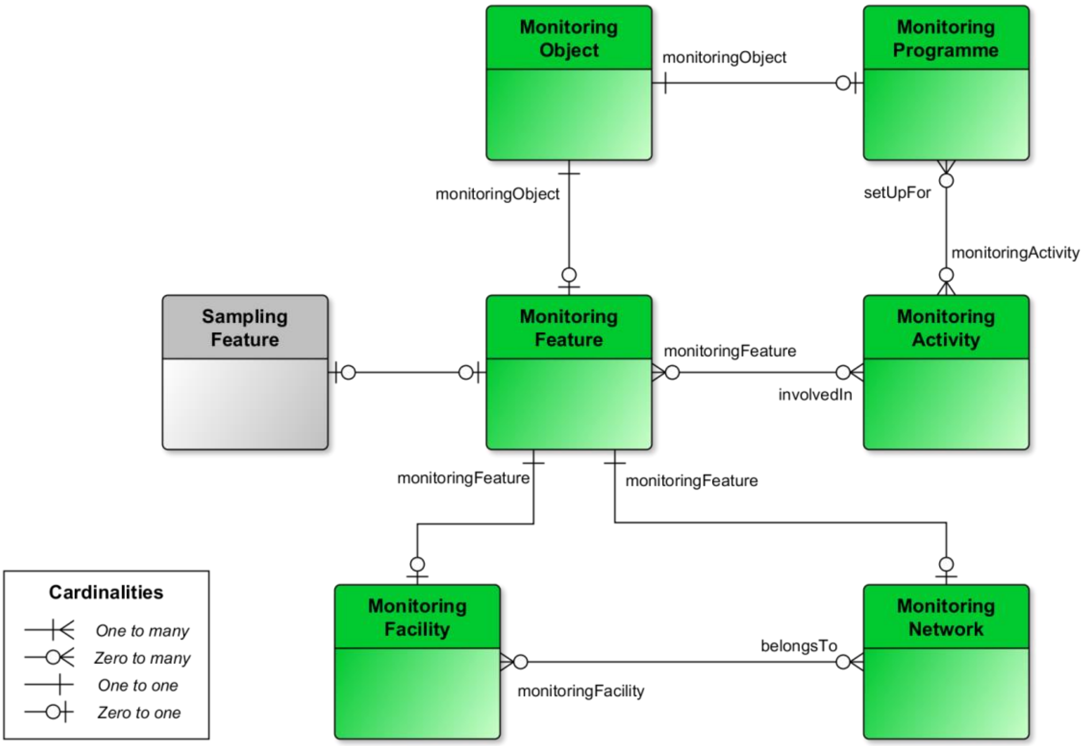
Some facilities may belong to both the Geology/Geophysics and Environmental Monitoring domains. In such cases while domain specific information is encoded both ways, observation metadata and results are published through the same Sampling Feature. For example, a water level monitoring station can be described as a Monitoring Facility, and/or a Borehole. However, measurement details and results would not be duplicated, because both are linked to one single Sampling Feature.

Table 3.24: Environmental Monitoring CDM – Entity definitions

Entity	Definition
Monitoring Object	An abstract base class for environmental monitoring objects. *Note:* It is a parent entity for all Monitoring Features and Monitoring Programme entities.
Monitoring Feature	An abstract base class for environmental monitoring features in the real world. *Note:* It is a parent entity for specialised features such as Monitoring Network and Monitoring Facility entities.
Monitoring Programme	Framework based on policy relevant documents defining the target of a collection of observations and/or the deployment of Monitoring Features on the field.
Monitoring Activity	Specific set of Monitoring Features used for a given domain. It is a concrete realisation of a given Monitoring Programme.
Monitoring Facility	A georeferenced object directly collecting or processing data about objects whose properties (e.g. physical, chemical, biological or other aspects of environmental conditions) are repeatedly observed or measured.
Monitoring Network	Administrative or organisational grouping of Monitoring Facilities managed the same way for a specific purpose, targeting a specific area.

Source: NEA, 2019.

Figure 3.3: Environmental Monitoring CDM – Entity Relationship Diagram



Source: NEA, 2019.

Table 3.25: Attributes for "Monitoring Object" entity

Entity	Attribute	Cardinality	Data Type	Description
Monitoring Object	**identifier**	[1..1]	*Identifier*	Unique Identifier.
	mediaMonitored	[1..1]	*GenericName*	Monitored environmental medium (e.g. air, biota, landscape, sediment, soil/ground, waste, water).
	geometry	[1..1]	*Geometry*	Geometry associated to the Monitoring Object.
	name	[1..1]	*String*	Name of Monitoring Object.
	additionalDescription	[1..1]	*String*	Plain text description of additional information not fitting in other attributes.
	responsibleParty	[1..1]	*"Responsible Party"*	Key responsible party for the Monitoring Object. Reference to the attributes of the "Responsible Party" entity of the O&M CDM (see section 4.4.7 of "RepMet Tools and Guidelines" [NEA, 2021c]).
	onlineResource	[1..1]	*URL*	A link to an external document providing further information on the Monitoring Object.
	purpose	[1..1]	*GenericName*	Reason for which the Monitoring Object has been set up.

Source: NEA, 2019.

Table 3.26: Attributes for "Monitoring Feature" entity[12]

Entity	Attribute	Cardinality	Data Type	Description
Monitoring Feature	***monitoringObject***	[1..1]	*"Monitoring Object"*	Reference to the parent "Monitoring Object" entity (see Table 3.25).
	involvedIn	[0..N]	*"Monitoring Activity"*	Reference to the "Monitoring Activity" in which the Monitoring Feature is involved (see Table 3.29).

Source: NEA, 2019.

Table 3.27: Attributes for "Monitoring Facility" entity

Entity	Attribute	Cardinality	Data Type	Description
Monitoring Facility	***monitoringFeature***	[1..1]	*"Monitoring Feature"*	Reference to the parent "Monitoring Feature" entity (see Table 3.26).
	representativePoint	[1..1]	*Geometry*	Representative location for the Monitoring Facility.
	measurementRegime	[1..1]	*GenericName*	Categories for different types of the measurement regime, e.g. continuous, demandDriven, onceOff, periodic.
	mobile	[1..1]	*Boolean*	Indicate whether the Monitoring Facility is mobile (repositionable) during the acquisition of the observation.
	resultAcquisition Source	[0..N]	*GenericName*	Source of result acquisition, e.g. exSitu, inSitu, remote, subsumed.
	specialisedEMFType	[0..N]	*GenericName*	Categorisation of Monitoring Facilities generally used by domain and in national settings, e.g. platform, sensor, site, station.
	belongsTo	[0..N]	*"Monitoring Network"*	Reference to the "Monitoring Network" which the Monitoring Feature belongs to (see Table 3.28).

Source: NEA, 2019.

12. "Monitoring Feature" is an abstract entity. It acts as a link between Monitoring Activity, Monitoring Facility, and Monitoring Network.

Table 3.28: Attributes for "Monitoring Network" entity

Entity	Attribute	Cardinality	Data Type	Description
Monitoring Network	***monitoringFeature***	[1..1]	*"Monitoring Feature"*	Reference to the parent "Monitoring Feature" entity (see Table 3.26).
	organisationLevel	[1..1]	*Geometry*	Level of legal organisation the monitoring network is affiliated with e.g. European, international, national, sub-national.
	monitoringFacility	[0..N]	*"Monitoring Facility"*	Reference to the "Monitoring Facility" entity (see Table 3.27).

Source: NEA, 2019.

Table 3.29: Attributes for "Monitoring Activity" entity

Entity	Attribute	Cardinality	Data Type	Description
Monitoring Activity	***monitoringFeature***	[0..N]	*"Monitoring Feature"*	Reference to the "Monitoring Feature" entity (see Table 3.26) that is involved in the activity.
	activityTime	[1..1]	*TimeRange*	Lifespan of monitoring activity.
	activityConditions	[1..1]	*String*	Textual description of monitoring activity.
	relatedParty	[1..1]	*"Responsible Party"*	Reference to the organisation responsible for the activity. Reference to the attributes of the "Responsible Party" entity of the O&M CDM (*See section 4.4.7 of* "RepMet Tools and Guidelines" [NEA, 2021c]).
	setUpFor	[0..N]	*"Monitoring Programme"*	Reference to related "Monitoring Programme" entities (see Table 3.30).
	onlineResource	[0..N]	*URL*	URL of external document providing further information.

Source: NEA, 2019.

Table 3.30: Attributes for "Monitoring Programme" entity

Entity	Attribute	Cardinality	Data Type	Description
Monitoring Programme	***monitoringObject***	[1..1]	*"Monitoring Object"*	Reference to the parent "Monitoring Object" entity (see Table 3.25).
	monitoringActivity	[0..N]	*"Monitoring Activity"*	Reference to the related "Monitoring Activity" entity (see Table 3.29).
	identifier	[1..1]	*Identifier*	Unique identifier for the monitoring programme.

Source: NEA, 2019.

Interconnection with the Repository Library

There is a strong connection between the Repository Library CDM (NEA, 2021b) and the Environmental Monitoring CDM in the Site Characterisation Library. Following the general data modelling rules, it is recommended that the "Repository Monitoring System" entity (related to the Repository Library) is "sub-classed" from the "Monitoring Feature" entity (related to the Site Characterisation Library) – meaning that a repository monitoring system has to be considered as a special kind of monitoring facility, with its attributes being specific controlled dictionaries dealing with the monitoring systems for the radioactive waste, or the other parameters of the repository engineered barrier system (R-EBS).

4. Example of applications of the CDMs

From a general point of view, a CDM provides a schema describing the structure of a database, with a CDM instance being the application of that schema for a real-world object such as a geological map or an abstract object such as an observation.

This chapter provides examples of how the CDMs presented in Chapter 3 can be applied in practice through presenting specific instances of CDMs for real-world objects. The geophysical examples also include the instances of the CDMs for the O&M standard and the MRMS that are introduced and explained in “RepMet Tools and Guidelines” (NEA, 2021c). The geological examples do not directly show the connection with the O&M CDM since it is implied in the attribute samplingFrame. The monitoring example contains a link to an Observation, but the full example is not provided: apart from property and process names this would be almost identical to the geophysical example.

4.1. Example application of Geology CDM

4.1.1. Geologic Unit – Geologic map

A geologic map is the representation of the geologic features of a certain site. It usually shows the different geologic units (i.e. volume of rocks with distinct characteristics) through different colours polygons and symbols.

Table 4.1 shows how the data related to each specific outcrop of the Dachstein Limestone formation can be structured as a Mapped Geological Unit based on the RepMet Geology CDM.

Table 4.1: Geology CDM applied to a "Mapped Geologic Unit"

<table>
<tr><th colspan="4">Entities</th><th>Attributes</th><th>Attribute values
(for illustrative purpose)</th></tr>
<tr><td rowspan="10">Mapped Feature</td><td rowspan="8">Geologic Unit</td><td rowspan="5">Geologic Feature</td><td rowspan="2"></td><td>identifier</td><td>GU_dT3.123</td></tr>
<tr><td>name</td><td>Dachstein limestone</td></tr>
<tr><td rowspan="3">Geologic Event</td><td>eventProcess</td><td>sedimentaryProcess</td></tr>
<tr><td>olderNamedAge</td><td>upper triassic</td></tr>
<tr><td>youngerNamedAge</td><td>upper triassic</td></tr>
<tr><td colspan="2" rowspan="2">Composition part</td><td>material</td><td>limestone</td></tr>
<tr><td>role</td><td>only part</td></tr>
<tr><td colspan="2"></td><td>unitType</td><td>lithologic unit</td></tr>
<tr><td colspan="3" rowspan="2"></td><td>shape</td><td>POLYGON(())</td></tr>
<tr><td>samplingFrame</td><td>SF_dT3.123.1</td></tr>
</table>

Source: NEA, 2019.

In this example, the Sampling Frame can provide a link across to an Observation. However, it is not usual practice to provide Observation data alongside geological maps, and this is not supported by the INSPIRE schema. It is also possible to add a link to a Sampling Feature via the Sampling Frame.

4.1.2. Geologic structure – mapped fault

A mapped fault is represented on a geologic map as a linear feature. Table 4.2 shows the encoding of this structure as a Mapped Fault.

Table 4.2: Geology CDM applied to a "Mapped Fault"

<table>
<tr><th colspan="4">Entities</th><th>Attributes</th><th>Attribute values
(for illustrative purpose)</th></tr>
<tr><td rowspan="8">Mapped Feature</td><td rowspan="6">Shear Displacement Structure</td><td rowspan="5">Geologic Feature</td><td rowspan="2"></td><td>identifier</td><td>SDS_Fault.123</td></tr>
<tr><td>name</td><td>Sierra Madre Fault</td></tr>
<tr><td rowspan="3">Geologic Event</td><td>eventProcess</td><td>subduction</td></tr>
<tr><td>olderNamedAge</td><td>miocene</td></tr>
<tr><td>youngerNamedAge</td><td>holcene</td></tr>
<tr><td colspan="2"></td><td>faultType</td><td>thrustFault</td></tr>
<tr><td colspan="3" rowspan="2"></td><td>shape</td><td>LINESTRING()</td></tr>
<tr><td>samplingFrame</td><td>SF_Fault.123.1</td></tr>
</table>

Source: NEA, 2019.

4.2. Example application of Geophysics CDM

4.2.1. Geophysical measurement - borehole logging measurement

Borehole logging is the process of in situ measurement of physical, chemical and structural properties of penetrated geological formations using sensors that are lowered into a borehole usually on a wireline cable. The borehole logging measurements are recorded in digital format as a function of depth (Wonik and Olea, 2017:431-474).

Data coming from borehole logging measurements can be structured using the RepMet Geophysics CDM. The borehole logging measurements are encoded through the "Geophysics Profile" entity as explained in Section 3.2.2, with the attribute "projected geometry" providing the location of the borehole.

Tables 4.2 to 4.6 show how the Geophysics CDM combined with the O&M CDM can be used to encode a 3D borehole logging measurement. Colour coding is used to help in the visualisation of the connection between the tables.

Table 4.3: Geophysics CDM applied to "Borehole Logging Measurement"

Entities				Attributes	Attribute values (for illustrative purpose)
Geophysics Profile	Geophysics Measurement	Geophysics Object	Sampling Feature	identifier	BHL_ASD-123
				shape	LINESTRING()
				sampledFeature	borehole:ASD-123
				relatedObservation	OBS_XXX-0001
				projectedGeometry	POINT(432543 654433)
				largerWork	NULL
				verticalExtent	depth; minValue: 0 m; maxValue: 1200 m
				platformType	ground
				relatedNetwork	NULL
				profileType	boreholeLogging

Source: NEA, 2019.

Table 4.3 reports the borehole logging geometry that is documented in the shape attribute of the embedded Sampling Feature as a 3D curve.

Table 4.4: O&M CDM (Observation entity) applied to “Borehole Logging Observation”

Entity	Attributes	Attribute values (examples)
Observation	identifier	OBS_XXX-0001
	name	Borehole Data Acquisition ASD-123/1
	responsible	operator: Edoardo Amaldi
	responsible	processor: Franco Rasetti
	responsible	reviewer: Emilio Segré
	observedProperty[1]	density
	observedProperty[2]	resistivity
	observedProperty[3]	porosity
	phenomenonTime	2000-01-01T12:00:00:00.000
	resultTime	2000-02-01T12:00:00:00.000
	parameter	samplingInterval=0.1 m
	parameter	depthMin=0
	parameter	depthMax=321
	parameter	instrument=INST_LogTechPro-0012
	procedure	Borehole Data Acquisition: PRC_XXX-0001
	result	RSC_XXX-0001.1

Source: NEA, 2019.

Table 4.4 lists the details of the observation related to the borehole logging, referenced and encoded as *Observation OBS_XXX-0001* according to the O&M CDM.

Table 4.5: O&M CDM (Process entity) applied to “Borehole Logging Process”

Entity	Attributes	Attribute values (for illustrative purpose)
Process	identifier	PRC_XXX-0001
	name	Borehole Data Acquisition
	type	boreholeDataAcquisition
	documentation	http://repmet/processes/boreholeDataAcquisition.html
	processParameter[1]	samplingInterval
	processParameter[2]	depthMin
	processParameter[3]	depthMax
	responsibleParty	custodian: Bruno Pontecorvo

Source: NEA, 2019.

Table 4.5 explains the process used for the borehole logging observation, referenced and encoded as *Process PRC_XXX-0001* according to the O&M CDM.

Table 4.6: Resource reference to Borehole Logging Result

Entity	Attributes	Attribute values (for illustrative purpose)
Resource	**identifier**	RSC_XXX-0001.1
	url	not available
	fileName	//c:/BHL123/asd/123.las
	title	Borehhole Log data
	format	LAS 2.0

Source: NEA, 2019.

Table 4.6 details the availability of observation results – referenced and encoded as *Resource RSC_XXX-0001.1* according to the MRMS CDM. Resources can be found in a file system, on the web via a URL, or both. The example shows a digital copy that is only available as a file on a local file system.

4.2.2. Solid model - seismic volume

A Seismic Volume is a three-dimensional (3D) grid of acoustic properties reconstructed from 3D seismic measurements. The speed of acoustic waves depends on the underlying geology and a 3D seismic reconstruction can give a very detailed 'image' of the underground conditions. Seismic Volumes are encoded as a Solid Model. The Projected Geometry is a two-dimensional polygon outlining the exploration area, and it can be used for data discovery purposes.

The 3D bounding shell is stored in the shape attribute of the embedded Sampling Feature (Table 4.7).

Table 4.7: Geophysics CDM applied to "Seismic Volume"

Entity				Attribute	Attribute value (for illustrative purpose)
Solid Model	Geophysics Model	Geophysics Object	Sampling Feature	**identifier**	SGM_ASD-123
				shape	srsDimensions=3;MULTIPOLYGON((()))
				sampledFeature	SeismicSwath;S3D_ASD-123
				relatedObservation[1]	OBS_S3D_ASD-123.1 (timeMigration)
				relatedObservation[2]	OBS_S3D_ASD-123.2)depthMigration)
				projectedGeometry	POLYGON(())
				largerWork	NULL
				verticalExtent	depth; minValue: 0 m; maxValue: 5000 m
				modelType	SeismicVolume

Source: NEA, 2019.

Table 4.8: O&M CDM (Observation entity) applied to "3D Seismic Depth Migration Observation"

Entity	Attribute	Attribute value (for illustrative purpose)
Observation	**identifier**	OBS_S3D_ASD-123.2
	name	3D Seismic Depth Migration ASD-123
	responsible	operator: Oscar D'Agostino
	responsible	processor: Ettore Majorana
	responsible	reviewer: Enrico Fermi
	observedProperty	Seismic Amplitude
	phenomenonTime	2000-01-01T12:00:00:00.000
	resultTime	2000-02-01T12:00:00:00.000
	parameter[1]	algorithm=kirchhoff
	parameter[2]	resolutionX=1 m
	parameter[3]	resolutionZ=1 m
	parameter[4]	processingSupportFile=ASD-123.2.sps
	procedure	3D Depth migration: PRC_3DDPTH_MIG
	result	RSC_S3D_ASD-123.2.1

Source: NEA, 2019.

Table 4.8 lists the details of one of the two observations (i.e. the 3D seismic depth migration) related to the seismic volume referenced and encoded as *Observation OBS_S3D_ASD-123.2* according to the O&M CDM.

Table 4.9: O&M CDM (Process entity) applied to the "3D Seismic Depth Migration Process"

Entity	Attribute	Attribute value (for illustrative purpose)
Process	**identifier**	PRC_3DDPTH_MIG
	name	3D Depth Migration
	type	3DSeismicDepthMigration
	documentation	http://repmet/processes/3DSeismicDepthMigration.html
	processParameter[1]	algorithm
	processParameter[2]	resolutionX
	processParameter[3]	resolutionZ
	processParameter[4]	processingSupportFile
	responsibleParty	custodian: Enrico Persico

Source: NEA, 2019.

Table 4.9 explains the process (i.e. the 3D depth migration) used for the above observation, referenced and encoded as *Process PRC_3DDPTH_MIG* according to the O&M CDM.

Table 4.10: Resource reference to 3D Seismic Depth Migration Result

Entity	Attribute	Attrivute value (for illustrative purpose)
Resource	**identifier**	RSC_S3D_ASD-123.2.1
	url	http://any.company/3dseismics/asd-123.pdf
	fileName	//c:/BHL123/asd/123.pdf
	title	3D Seismic Cube from Depth Migration
	format	PDF

Source: NEA, 2019.

Table 4.10 details the availability of observation results, referenced and encoded as *Resource RSC_S3D_ASD-123.2.1* according to the MRMS CDM. Resources can be found in a file system, on the web via a URL, or both. The example shows a digital copy that is only available as a file on a local file system.

4.3. Example application of Environmental Monitoring CDM

4.3.1. Monitoring Facility – air pollution monitoring station

In Table 4.11, an air pollution monitoring station is encoded as a Monitoring Facility. General metadata attributes like name, description, responsible party are contained in the Monitoring Object part. More specific attributes characteristic to the technical implementation are associated with the Monitoring Facility. There is also a link to a Sampling Feature, so optionally it has shape and Observation. OBS_XXX-1001.1 is the identifier of the related Observation that contains the air pollution data recorded by the monitoring station.

Table 4.11: Data Model applied to Monitoring Facility

Entity			Attribute	Value (for illustrative purpose)
Monitoring Facility	Monitoring Feature	MonitoringObject	**mediaMonitored**	air
			name	Veszprem I.
			additional Description	Example Monitoring Station
			responsibleParty	custodian: Nella Mortara
			onlineResource	http://example.org/MF/info.html
			purpose	Air pollution monitoring
		Sampling Feature	**identifier**	MF_XXX-1001
			shape	POINT(47.09 17,9)
			sampledFeature	http://sweetontology.net/matr/Air
			relatedObservation	OBS_XXX-1001.1
			representativePoint	POINT(47.09 17,9)
			measurementRegime	continuous
			mobile	false
			resultAcquisition Source	inSitu
			specialisedEMFType	station

Note: The online resources URL in this table is provided as an example only and does not currently link to any live resource.

Source: NEA, 2019.

5. Controlled dictionaries

5.1. Introduction to controlled dictionaries and their place in RepMet

Controlled dictionaries (also called controlled vocabularies) are collections of agreed terms that a community or an organisation uses, manages and maintains in a controlled way within a particular domain of interest. They play a fundamental role in harmonisation of data and information systems, supporting system interoperability and long-term usability. On the data provider side, controlled dictionaries help the development of uniform content, whereas, on the data user side, they support queries and understanding. Modern controlled dictionaries are often implemented using the technologies and standards of the World Wide Web, such as the international standards that the World Web Consortium (W3C) has developed. The three RepMet Libraries include web-based controlled dictionaries developed with the RDF/SKOS standard originating with the W3C. Chapter 3 of the "RepMet Tools and Guidelines" report (NEA, 2021c) provides an introduction to controlled dictionaries, why they are useful, and the technical bases underlying them, with examples from the domain of RepMet.

5.2. Controlled dictionaries in the Site Characterisation Library

RepMet adopted the controlled dictionaries that the INSPIRE community has already made available online for all the 34 spatial data themes in the INSPIRE framework: they are directly accessible on the internet at the central "INSPIRE Code List Register". The INSPIRE Code List Register provides an initial set of dictionaries, harmonised across different areas of geoscience, that are intended to be extended in the future.

For the Site Characterisation Library, these code lists represent controlled dictionaries for the attributes of the entities of the Geology, Geophysics and Environmental Monitoring CDMs in Chapter 3. The code lists are hierarchical controlled dictionaries defining the values that the attributes can assume: for example, the "olderNameAge" attribute of the "Geologic Event" entity of the Geology CDM can take values such as "Aalenian", "Aeronian", etc. Each controlled dictionary is referenced by a specific Uniform Resource Locator (URL). At this URL, the controlled dictionaries are available in several formats including RDF/SKOS. The tables in the following paragraphs specify the controlled dictionary URL for each attribute.

In future, there is scope for RepMet to contribute to the INSPIRE work by formulating list of attributes for the monitoring activities that are carried out in the radioactive waste management facilities such as final repositories.

Box 5.1: Controlled dictionaries: Site Characterisation vs Waste Package and Repository Libraries

The RepMet group followed two different approaches for the development of the controlled dictionaries in the Site Characterisation and the Waste Package / Repository Libraries. For the Site Characterisation, RepMet was able to take advantage of the controlled dictionaries that the INSPIRE community had already created at that time. This approach avoids "reinventing the wheel" through building on the work already done in INSPIRE. By contrast, for the Waste Package/Repository Library, the RepMet group produced its own RDF/SKOS controlled dictionaries describing the two library topics, i.e. "packaged waste and spent nuclear fuel ready for final disposal at the repository" and "repository requirements and structures at closure", respectively. No existing controlled dictionaries, at least at international or similar scale, about these two topics were available. The Waste Package Library and the Repository Library are original RepMet outputs that follow international standards such as the W3C RDF/SKOS, and constitute an important contribution for the radioactive waste management communities.

5.2.1. Geology

Table 5.1 shows the controlled dictionary for the attributes of the Geology CDM entities. The listed attributes should take values only from the controlled dictionary located at the corresponding URL.

For the Geology domain case, RepMet considered not only the INSPIRE controlled dictionaries, but also those of the CGI (Commission for the Management and Application of Geoscience Information), which provide an extensive set of controlled dictionaries for GeoSciML available on the web.[13]

13. http://resource.geosciml.org/vocabulary/cgi/

Table 5.1: Controlled dictionaries (code lists) for Geology CDM[14]

Entity	Attribute	Organisation	URL[15]
Geologic Event	eventEnvironment	INSPIRE	http://inspire.ec.europa.eu/codelist/EventEnvironmentValue
	eventProcess	INSPIRE	http://inspire.ec.europa.eu/codelist/EventProcessValue
	olderNamedAge	INSPIRE	http://inspire.ec.europa.eu/codelist/GeochronologicEraValue
	youngerNamedAge	INSPIRE	
Shear Displacement Structure	faultType	INSPIRE	http://inspire.ec.europa.eu/codelist/FaultTypeValue
Fold	profileType	INSPIRE	http://inspire.ec.europa.eu/codelist/FoldProfileTypeValue
Contact	contactType	CGI	http://resource.geosciml.org/vocabulary/cgi/201211/contacttype.html
Geologic Unit	geologicUnitType	INSPIRE	http://inspire.ec.europa.eu/codelist/GeologicUnitTypeValue
Composition Part	material	INSPIRE	http://inspire.ec.europa.eu/codelist/LithologyValue
	role	INSPIRE	http://inspire.ec.europa.eu/codelist/CompositionPartRoleValue

Source: NEA, 2019.

5.2.2. Geophysics

Controlled dictionaries (or code lists) for the entity attributes of the Geophysics CDM are listed in Table 5.2. Except for those defined for measurement and model types, these code lists are in an early state of development and they will be further developed in future by the INSPIRE community. More dictionaries are available in the extension model that can be found in the INSPIRE Technical Guidelines.

14. The applicability of the above controlled dictionaries depends on the map scale: for example, the collections of terms about "lithology" and "geologic age" are optimised only up to the 1:1 000 000 scale geology map of Europe. For maps on a more detailed regional scale the integration of national or regional code lists is essential, but will require great efforts from the community. To support this future activity, INSPIRE code lists are extendable.

15. All websites accessed June 2019.

Table 5.2: Controlled dictionaries (code lists) for Geophysics CDM

Entity	Attribute	Organisation	URL[16]
Geophysics Measurement	platformType	INSPIRE	http://inspire.ec.europa.eu/codelist/PlatformTypeValue
Geophysics Station	stationType	INSPIRE	http://inspire.ec.europa.eu/codelist/StationTypeValue
	stationRank	INSPIRE	http://inspire.ec.europa.eu/codelist/StationRankValue
Geophysics Profile	profileType	INSPIRE	http://inspire.ec.europa.eu/codelist/ProfileTypeValue
Geophysics Swath	swathType	INSPIRE	http://inspire.ec.europa.eu/codelist/SwathTypeValue
Curve Model	modelType	INSPIRE Technical Guidelines	http://inspire.ec.europa.eu/codelist/CurveModelTypeValue
Surface Model	modelType	INSPIRE Technical Guidelines	http://inspire.ec.europa.eu/codelist/SurfaceGridModelTypeValue
Solid Model	modelType	INSPIRE Technical Guidelines	http://inspire.ec.europa.eu/codelist/SolidGridModelTypeValue
Campaign	campaignType	INSPIRE	http://inspire.ec.europa.eu/codelist/CampaignTypeValue
	surveyType	INSPIRE	http://inspire.ec.europa.eu/codelist/SurveyTypeValue

Source: NEA, 2019.

5.2.3. Environmental Monitoring

Basic controlled dictionaries (or code lists) for the entity attributes of the Environmental Monitoring CDM are listed in Table 5.3. As for the geophysics domain, these controlled dictionaries are in an early stage of development and they will be extended in future projects within the INSPIRE community.

Table 5.3: Controlled dictionary (code lists) for Environmental Monitoring CDM

Entity	Attribute	Organisation	URL[17]
Monitoring Object	mediaMonitored	INSPIRE	http://inspire.ec.europa.eu/codelist/MediaValue
	purpose	INSPIRE	http://inspire.ec.europa.eu/codelist/PurposeOfCollectionValue
Monitoring Facility	measurementRegime	INSPIRE	http://inspire.ec.europa.eu/codelist/MeasurementRegimeValue
	resultAcquisitionSource	INSPIRE	http://inspire.ec.europa.eu/codelist/ResultAcquisitionSourceValue
	specialisedEMFType	INSPIRE	http://inspire.ec.europa.eu/codelist/SpecialisedEMFTypeValue
Monitoring Network	organisationLevel	INSPIRE	http://inspire.ec.europa.eu/codelist/LegislationLevelValue

Source: NEA, 2019.

16. All websites accessed June 2019.

17. All websites accessed June 2019.

6. Concluding remarks

The Site Characterisation Library is a technical product of the Nuclear Energy Agency (NEA) Integration Group for the Safety Case (IGSC) RepMet initiative. It is composed of a report (this document) and an associated technical library dealing with data and related metadata that are considered during the characterisation of a site investigated and surveyed for suitability for radioactive waste disposal purposes, leading up to site selection.

The "Site Characterisation Library" is a technical report designed to show the application of metadata tools and techniques within the area of geological and geophysical characterisation of the repository site for radioactive waste management. Rather than developing new standards, the RepMet group reviewed a range of existing national and international geoscience data standards and selected a number that can be usefully applied in the area of radioactive waste management.

The Site Characterisation Library has two principal aims:

- to show how the use of appropriate metadata can support the long-term management of the "core information" that is acquired during the characterisation of a site investigated and surveyed for suitability for radioactive waste disposal purposes, leading up to site selection;
- to provide applied examples of how implementing metadata-based techniques can support the long-term management of the "core information".

The library includes high-level conceptual data models, descriptions of data entities, attributes, associated metadata and controlled dictionaries. The library also includes application examples from geology, geophysics and environmental monitoring. Radioactive Waste Management Organisations (RWMOs) can reuse and further extend the models and controlled dictionaries in the development of their own data and information systems.

Prior to the establishment of the RepMet initiative there was a lack of national and international metadata standards that specifically supported the management of radioactive waste. Therefore, the RepMet group reviewed a range of metadata standards and selected a number of them that, even if originally not related or designed for this area, are based on generic concepts and schemas that can be easily adapted and applied to this field. The Site Characterisation Library relies heavily on the data models developed as part of the INSPIRE (Infrastructure for Spatial Information in Europe) initiative (European Union, 2007). The other selected standards are O&M, MRMS and the W3C RDF/SKOS.

Although not created specifically for radioactive waste management activities, one major topic of the INSPIRE initiative is geodata, and INSPIRE data models provide a substantial element of the information required for site characterisation. RepMet made additions and modifications in some areas in order to allow development of the conceptual high-level data model for site characterisation. A more in-depth examination of INSPIRE in terms of completeness for radioactive waste management projects was not possible within the time available. This should be done as part of the development of logical data models.

With the Site Characterisation Library, RepMet has taken the first step towards harmonising the site characterisation activities of RWMOs. Further steps are now needed to help facilitate the transferability, verifiability and transparency of activities and of the underlying data.

Although the RepMet initiative has now finished, there is further work that can be done. This includes the improvement of the controlled dictionaries included in the Site Characterisation Library. The controlled dictionaries could then become an international resource curated by the NEA.

Other activities include:

- Further development of the scientific and technical content of the controlled dictionaries (e.g. more details for "definition" and "purpose" features for each attribute).
- Definition of a strong connection between the attributes of the controlled dictionaries and the NEA International Features, Events and Processes (IFEP) List included in the NEA FEP Database. This is because each item of the NEA IFEP List reports and explains their eventual relevance for safety assessment.
- Elaboration of controlled dictionaries for attributes of entities in the O&M and MRMS standards.

References

Andersson, J. (2003), "Site descriptive modelling – strategy for integrated evaluation", SKB Sweden.

CGI. (n.d.). GeoSciML Data Model, retrieved from www.geosciml.org/geosciml/3.2/documentation/html/ (accessed June 2019).

Cox, S. (ed.), Open Geospatial Consortium Inc. (2013). Geographic Information: Observations and Measurements, OGC Abstract Specification Topic 20, OGC 10-004r3 and ISO/DIS 19156, retrieved from www.ogc.org/standards/om (accessed June 2019).

DCMI Usage Board, "Dublin Core Metadata Initiative", retrieved from http://dublincore.org (accessed June 2019).

EU (2007, April 25). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Official Journal of the European Union. Retrieved from https://inspire.ec.europa.eu/ (accessed June 2019).

INSPIRE Thematic Working Group Environmental Monitoring Facilities (2013), D2.8.II/III.7 Data Specification on Environmental Monitoring Facilities – Technical Guidelines, retrieved from http://inspire.ec.europa.eu/file/1535/download?token=nbWSNl_1 (accessed June 2019)

INSPIRE Thematic Working Group Geology (2013), D2.8.II.4 INSPIRE Data Specification on Geology – Technical Guidelines, Retrieved from http://inspire.ec.europa.eu/file/1519/download?token=IGCGbum3 (accessed June 2019).

NEA (2021a), "Waste Package Library", OECD Publishing, Paris.

NEA (2021b), "Repository Library", OECD Publishing, Paris.

NEA (2021c), "RepMet Tools and Guidelines", OECD Publishing, Paris.

NEA (2018), *Metadata for Radioactive Waste Management*, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_15062/metadata-for-radioactive-waste-management.

Open Geospatial Consortium Inc. (n.d.), "Geographic information – Metadata", retrieved from www.iso.org/obp/ui/#iso:std:iso:19115:-1:ed-1:v1:en (accessed June 2019).

Open Geospatial Consortium (n.d.), SensorML: Model and XML Encoding Standard, retrieved from https://portal.opengeospatial.org/files/?artifact_id=55939 (accessed June 2019).

Open Geospatial Consortium (2007a), Observations and Measurements – Part 2 - Sampling Features, retrieved from www.opengeospatial.org/standards/om#downloads (accessed June 2019).

Portele, C. (ed.), Open Geospatial Consortium Inc. (2007b), OpenGIS Geography Markup Language (GML) Encoding Standard, OGC 07-036, Version 3.2.1, retrieved from http://portal.opengeospatial.org/files/?artifact_id=20509 (accessed June 2019).

Recordkeeping Metadata Development Committee (RMDC) (2015), Minnesota Recordkeeping Metadata Standard, retrieved from www.mnhs.org/preserve/records/metamrms.php (accessed June 2019).

Wonik, T. and R. A.Olea (2017), "Borehole Logging", In Knödel, K., G. Lange, and V. Voigt (2007), Environmental Geology: Handbook of Field Methods and Case Studies, Springer, Berlin, pp. 431-474.

Annex A. INSPIRE and Site Characterisation Library

There are six schemas in the INSPIRE data model that are most relevant to Site Characterisation.

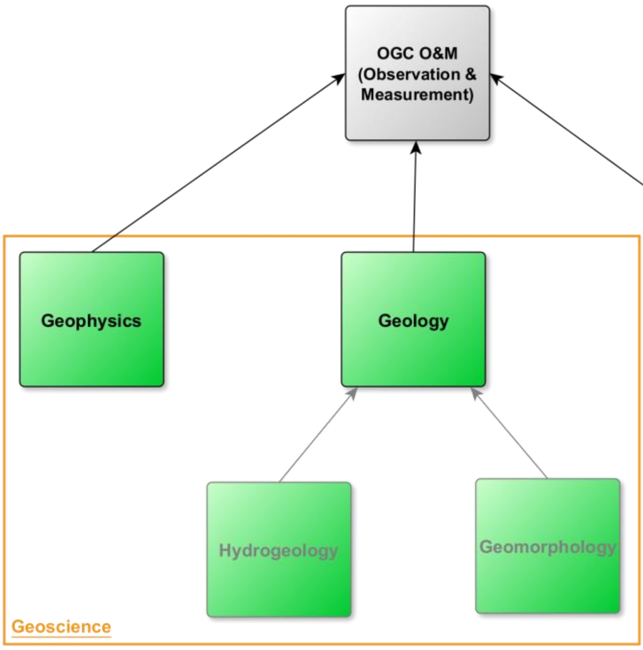
Figure A1 shows the scientific domains and their relations as they are represented in the existing standard. These are *geology*, *hydrogeology*, *geomorphology*, *geophysics*, *environmental monitoring* and *sampling*.

Geology, hydrogeology, geomorphology and geophysics belong together in the INSPIRE geology spatial data theme. Environmental Monitoring is defined in a separate schema. Monitoring facilities often collect information about geological and geophysical phenomena as well.

Scientific domains that have been separated for historical, technical and legal reasons are also separated in the INSPIRE data model. These data also have their importance for site characterisation, repository design and safety assessment, but they will not be handled here.

Data specification was undertaken by separate thematic working groups with specific background knowledge and requirements from different communities. During the development, harmonisation and interoperability were among the main issues. Special care was taken by cross-thematic groups to find similarities between different data themes in order to identify common elements and, if possible, use them as common building blocks. After the INSPIRE model had been consolidated no further harmonisation was undertaken. However, this work could be continued during the implementation phase.

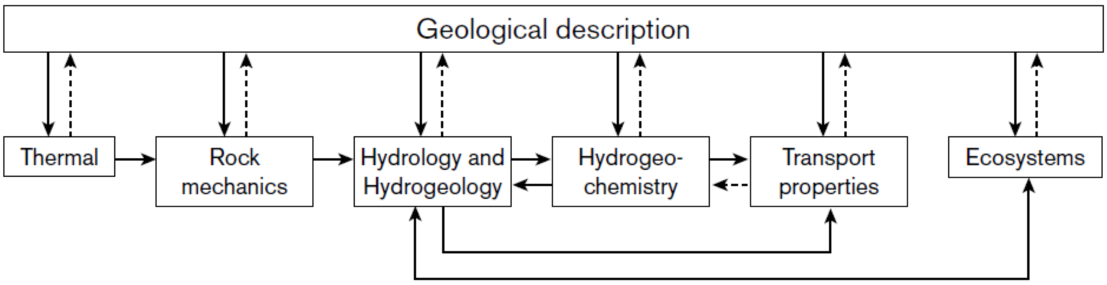
Figure A1: Geoscience domains as defined in INSPIRE



Source: NEA, 2019.

One of the common building blocks is the Open Geospatial Consortium (OGC) Observation and Measurements (O&M) standard (ISO19156) which creates an important bridge between seemingly separate domains. The O&M concept allows geosciences and environmental monitoring to be integrated into one generic data model. Domain specific information is described in the thematic packages, but in the centre there is the "Sampling Feature", a common entity to describe all observations in the same way.

Figure A2: Different discipline descriptions and the geologic framework (J. Andersson)



Source: Andersson J., 2003.

Figure A2 shows a process chain of different disciplines identified by J. Andersson in 2003 (Andersson, 2003) that are used to determine a complete 3D model of the geological and physical environment. It involves a large number of observations of different properties, carried out using different techniques. Geology provides the framework for the investigations, and is also the ultimate feature of interest for all observations. The geological description is continuously developing as new results are integrated into the model. The procedure can further be improved by adding more disciplines that are missing from the processing chain (e.g. mineralogy, geophysics or satellite imaging). There is a strong relation between Figure A1 and Figure A2 Schema packages provided by INSPIRE, and the related standards, support encoding and storing information collected in the procedure of geo-scientific investigations. This means that the outlined set of standards cover most of what is required for site characterisation.