

**OECD NEA Radioactive Waste Repository Metadata Management (RepMet) Initiative (2014-2018)
– 15614**

Claudio Pescatore*, Alexander Carter**

*OECD Nuclear Energy Agency¹ (claudio.pescatore@oecd.org)

** Radioactive Waste Management Limited, Nuclear Decommissioning Authority
(alexander.carter@nda.gov.uk)

ABSTRACT

The OECD/NEA Radioactive Waste Repository Metadata Management (RepMet) initiative aims to bring about a better understanding of the identification and administration of metadata – a key aspect of data management – to support national programmes in managing their radioactive waste repository data, information and records in a way that is both harmonised internationally and suitable for long-term management and use. At this point in the project, the main focus lies on “waste packages at their storage sites ready for geological disposal” but excluding spent fuel and high level waste (HLW). Later work will deal with metadata collected on spent fuel and HLW as well as with different phases of repository development: design and construction, and operation. In order to develop metadata, RepMet members have provided lists of data items they currently collect and manage for packaged waste. The lists, albeit it non comprehensive, suggest that the number and type of data items varies importantly from organisation to organisation. At term, RepMet plans to generate a comprehensive list of such data items and their metadata. In parallel, RepMet is developing a document which examines a selected number of existing metadata standards. Because of the specialist nature of radioactive waste management, including waste processing facilities, disposal facilities and waste packages, existing metadata standards do not seem to be transferable as defined. Rather, RepMet will explore such standards to identify relevant principles for possible adaptation to the radioactive waste management community.

INTRODUCTION

National radioactive waste repository programmes are collecting large amounts of data on a large number of data items to support the long-term management of their nations' radioactive waste.

Metadata is data about an object, which can be a physical item such as a book, CD, DVD, map, chair, table, flower pot, etc. or an electronic record such as a digital image, digital photo, document, programme file, database table, etc. Metadata provides information about one or more aspects of an object, such as:

- Means of creation of the object;
- Purpose of creation;
- Time and date of creation;
- Creator or author;
- Location on a computer network where the electronic object is stored;
- Standards for creating the electronic object;
- Etc.

For instance, an ISBN constitutes an item of metadata about a book (i.e., a data item, an object and a type of record) without referring to its contents. Similarly, a modern digital image file often includes

¹ *The opinions expressed and arguments employed herein are those of the author and do not necessarily reflect the official views of the OECD or of the governments of its member countries.*

WM2015 Conference, March 15 – 19, 2015, Phoenix, Arizona, USA

embedded metadata to describe the type of camera used or exposure settings.

The OECD/NEA Radioactive Waste Repository Metadata Management (RepMet) initiative aims to bring about a better understanding of the identification and administration of geological disposal related data and its associated metadata – a key aspect of data management – to support national programmes in managing their radioactive waste repository data, information and records in a way that is both harmonised internationally and suitable for long-term management and use [1]. This initiative is complementary to the international initiative on the Preservation of Records, Knowledge and Memory across Generations (RK&M) [2], which focusses on the time after the closure of a repository.

At present, 12 organisations from 12 countries participate in the RepMet initiative, whose second meeting took place in September 2014. The next meeting will take place in Spring 2015.

BACKGROUND ON DATA AND METADATA

Before discussing the details of (meta-)data collection, it is worth highlighting that the boundary between data and metadata often depends on context and use. Bibliographic records associated with a book may be regarded as metadata by the person searching a particular book; the librarian, however, may regard it as data. The distinction between data and metadata made is, as a result, not clear-cut. However, a tentative distinction is useful for analytical purposes.

Data

Typically, a national radioactive waste management programme has several systems of data, information, and/or records management, including databases that serve the needs of different communities. Subjects traditionally covered include:

- Geoscience;
- Engineering;
- Waste management; and
- Context aspects.

Data is used for a variety of purposes including:

- Site characterisation and site selection;
- Numerical modelling;
- Repository design and construction;
- Repository operation;
- Repository licensing;
- Waste packaging and conditioning;
- Environmental impact assessment;
- Quality assurance; and
- Financial, political and legislative purposes.

Data and related records have to be collected and managed so that future generations of specialists both within and outside waste management organisations (WMOs) will still be able to make sense of them. Non-specialist audiences are also very much interested in radioactive waste repository information,

WM2015 Conference, March 15 – 19, 2015, Phoenix, Arizona, USA

especially information concerning radiation protection and environmental safety, and may need to retrace and understand decisions made in the past.

The relatively long time scales involved call for management practices for data and related records that are robust and, as far as possible, future-proof. Data must be preserved in a manner that continues to be intelligible, accessible and searchable in the future despite the inevitable changes in personnel and technology within organisations. Robustness also requires resiliency to future operational and societal changes, such as the advent of new organisational structures and evolutions in the decision making environment. In general it is important that:

- The right information has been captured;
- A medium which will last and can be read in future has been used (paper, copper, magnetic media or whatever);
- The information uses a format which can be understood in future (encoding, file formats);
- The collection contains an index and/or structure so that relevant information can be located when desired; and
- The collection is stored in one (or more) safe and secure locations.

Metadata

The main purpose of metadata is to facilitate the re-accessing of information relevant to the original object or which adds context to it in some way. In the case of RK&M and RepMet we are concerned with data and metadata about physical or electronic records that describe experimental measurements, physical processes, decisions made at a certain point in time, safety files, etc.

Metadata can be seen as data and can be operated upon as data. Like data, metadata can be stored in either human-readable or binary form. Storing data and metadata in a human-readable format, such as XML, can be useful because users can understand and edit it without specialised tools, which may not be available in the future. As with data, metadata may be stored in an electronic database system to facilitate searching and reporting. In a relational database each piece of data or metadata is stored as a row in a named table whose columns have a name, size and type. These, in turn, may be thought of as metadata about the database.

There exist international standards to help define metadata and to operate with them. These standards may vary from application to application. There thus exist: classes of metadata and standards for describing geographical objects (such as datasets, maps, features, or simply documents with a geospatial component); similarly, in the field of ecology and environment, metadata and metadata standards are intended to document the who, what, when, where, why, and how of data collection for a particular study; elsewhere, metadata can be used to name, describe, catalogue and indicate ownership or copyright for a digital audio file (CDs such as recordings of music will carry a layer of metadata about the recordings such as dates, artist, genre, copyright owner, etc. This metadata, often not displayed by CD players, can be accessed and displayed by specialized music playback and/or editing applications); etc.

REPMET METADATA LIBRARY

The RepMet survey regarding metadata on waste packages ready for disposal, discussed below, is a first attempt at developing a better understanding of data and its associated metadata in an area specific to the domain of radioactive waste management. In the context of the latter, context setting information may include information on:

WM2015 Conference, March 15 – 19, 2015, Phoenix, Arizona, USA

- Quality checking or approval;
- Provenance; and
- Ownership.

The final role of RepMet metadata is to ensure the long-term understandability and confidence in the data object, including by facilitating the exchange of information across borders and aiding international review. Encouraging the structuring of metadata in a common, internationally harmonised way – one of the overall objectives of RepMet – also has the purpose of ensuring that future generations can have access to and understand data collected in the present.

In order to develop an internationally harmonised metadata library it is important to understand which data items are currently managed in national waste management organisations. To do this the RepMet initiative developed a questionnaire about data items on “waste packages ready for disposal”.

The drafting of the questionnaire proved to be a difficult exercise requiring several reviews by participants. It became soon clear that the latter had different interpretations of the questions and did not have a consistent understanding of terminologies. It also became apparent during the process that the questions of safeguards, criticality and transportation, and to include HLW and spent fuel, will require additional specific attention in future.

The questionnaire covered the following five categories, and respondents were asked which data items were covered in each of the categories:

- Waste package receipt;
- Waste package physic-chemical properties;
- Waste container;
- Wasteform (i.e. waste + waste encapsulant); and
- Waste package handling.

A total of 11 organisations completed the questionnaire. Replies varied considerably, both in terms of number of data items and of additional items used. This spread in the number of items collected by WMOs was between 6 and 68. In fact, the only common item collected by all 11 respondents is the “unique waste package identifier”, but this may be down to terminology – at some level, organizations must be storing quantities of similar data.

Feedback from participants in the initiative on completing this questionnaire has provided interesting insights. For instance, one organisation indicated that, when completing the questionnaire, they used their Waste Acceptance Criteria (WAC). However, WACs appear to be different at different facilities of the same organisation, which made the completion of the questionnaire difficult. Other respondents pointed out that organisations often operate multiple databases and each may store a different set of information or data items. This bears the question of why such differences exist, how the different items are related to one another, and how traceability is ensured within organisations. In some cases such differences may be logical of course, for example a HLW database would probably be expected to contain different information to an ILW one. More likely to cause concern are historic differences which have evolved over time and may not have a good reason.

Discussing the use of metadata in any given field leads to the inevitable question about standards to be used. This not only applies to formats for metadata elements, but also to vocabularies used and the relationships between items. A controlled vocabulary provides a managed and maintained reference so that all parties understand terminology and metadata in the same way. The RepMet initiative has conducted a review of existing metadata standards that could be potentially relevant to the field of

WM2015 Conference, March 15 – 19, 2015, Phoenix, Arizona, USA

radioactive waste management. Based on this review, it will develop an annotated list of the following, potentially relevant metadata standards:

- INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119;
- Content Standard for Digital Geospatial Metadata;
- ISO 19115:2003 Geographic information – Metadata;
- ISO 19119:2005 Geographic information – Services;
- Directory Interchange Format;
- Dublin Core Scheme;
- ISO 23081 Information and documentation – Records management processes – Metadata for records; and
- MoReq2010.

The RepMet initiative will also agree a vocabulary (glossary). For the latter task it will start from the vocabulary created under the RK&M initiative [3].

Reference Workflow

The RepMet initiative is also developing a reference workflow for waste packages. This workflow chart would indicate decision points in the process of disposal at which organisations responsible for the package would take decisions. It would provide a logical visualisation of the lifecycle of a waste package.

In the context of RepMet, the decision points in the workflow would be associated with (meta-) data needed to make an informed decision. This would help in developing a shared understanding of the metadata used and needed at different points in the disposal process and foster consistency across and within organisations. The results of the questionnaire discussed above would be linked to different points in the reference workflow chart.

CONCLUSIONS

The OECD/NEA RepMet initiative examines data and metadata currently collected by waste management organisations in OECD countries to develop an understanding about differences and commonalities. An initial questionnaire, focusing on waste packages at site ready for disposal, has indicated that significant differences exist in the data items that are collected and managed. Sometimes data are managed through different databases, which can lead to data dispersion and potential differences in structure and definition.

Based on these results, the RepMet initiative is working to generate, in 2015-2016, a list of data items and its associated metadata for a Waste Package at a storage site ready for disposal. The final list will be synthesized from the lists supplied by members as currently used in their facilities. The aim is to generate a comprehensive list against which members may benchmark their practices. There is similarity with the NEA FEP list.

Beyond 2016 RepMet aims to help waste management organizations by developing a number of reference libraries for data and metadata in other areas related to geological disposal. This will assist organizations by providing consistency and completeness checks, and provide a harmonized structure, which organisations can adapt to their purposes. These metadata libraries will be linked to a reference workflow chart, the use of which will facilitate sharing and understanding the role of (meta-) data collected by WMOs.

REFERENCES

1. OECD/NEA, “Vision Document for the Radioactive Waste Repository Metadata Management (RepMet) Project”, 26 February 2014, available online at <http://www.oecd-nea.org/rwm/docs/2014/rwm2014-2.pdf>.
2. OECD/NEA, “Preservation of Records, Knowledge and Memory (RK&M) across Generations”, 28 October 2014, available online at <http://www.oecd-nea.org/rwm/rkm/>.
3. OECD/NEA, “Glossary of Terms. Preservation of Records, Knowledge and Memory (RK&M) across Generations. Draft – Definitions as of 3 March 2014”, 3 March 2014, available online at <http://www.oecd-nea.org/rwm/docs/2011/rwm2011-14-rev4.pdf>.