UMM-Collection Metadata Concepts in CMR Metadata Collections

Interim Collection Comparison Report for Big Earth Data Initiative

Metadata Source: CMR Metadata Collections

Metadata Dialect: ISO 19115-2

Evaluation Target: UMM-Collection metadata profile

The Unified Metadata Model Collection (UMM-Collection) profile describes documentation concepts that are considered important for collection level metadata. The profile is includes three documentation levels (Required, Recommended and Optional).

Metadata Collection

- Alaska Satellite Facility (ASF)
- Crustal Dynamics Data Information System (CDDIS)
- Global Hydrology Resource Center (GHRC)
- Goddard Earth Sciences Data and Information Center (GES_DISC)
- Level 1 and Atmosphere Archive and Distribution System (LAADS)
- Land, Atmosphere Near real-time Capability for EOS (LANCEMODIS)
- Land, Atmosphere Near real-time Capability for EOS (LANCEAMSR2)
- Langley Research Center (LARC)
- Langley Research Center (LARC_ASDC) Atmospheric Science Data Center
- Land Process DAAC - EOS Core System (LPDAAC_ECS)
- National Snow and Ice Data Center Version 0 (NSIDCV0)
- National Snow and Ice Data Center EOS Core System (NSIDC_ECS)
- Ocean Biology Processing Group (OBPG)

- Oak Ridge National Laboratory (ORNL)
- Physical Oceanography DAAC (PODAAC)
- Socioeconomic Data and Applications Center (SEDAC)
- U.S. Geological Survey Earth Resources Observation Systems (USGS_EROS)
- (AU_AADC)
- ESA
- EUMETSAT
- ISRO
- JAXA
- LM_FIRMS
- NOAA_NCEI
- USGS_LTA

Overview

We examined over 15,000 metadata records from 26 collections extracted from the Common Metadata Repository (CMR) during April 2016. The links below connect to tables in google sheets which provide the average number of occurrences of UMM-Collection elements in each of these collections. A value of 1 or more typically (although not necessarily) indicates that the element is included one or more times in each record in a collection. A value < 1.0 is typically the percentage of records in a collection that include the metadata element. Cells with pink backgrounds indicate values of 0, meaning the element is completely missing from the collection.

How are UMM-Collection Required concepts used in the CMR?

Overall CMR metadata collections are doing a very good job of documenting Required UMM-Collection concepts. The link below shows a google sheets comparison view of Required UMM-Collection concept usage across CMR metadata collections. The pink shaded cells near the bottom of the table indicate required elements that are missing for some of the collections.

Table 1: Collection Comparison of UMM-Collection Required Elements

Highlights

- 83% of CMR Collections include all required UMM-Collection concepts**.**
- An average of 20 science keywords are included in CMR metadata collections
- Spatial Extents are included in all CMR metadata collection containing more then one record.

Improvement Focus Areas

- The Project Name concept is missing from 4 CMR metadata collections**.**
- The Resource Version concept is missing from 8 NASA collections

How are UMM-Collection Recommended concepts used in the CMR?

The link below provides a google sheets comparison view of Recommended UMM-Collection concept usage across CMR metadata collections. The pink shaded cells near the bottom of the table indicate recommended elements that are missing for some of the collections.

Table 2: Collection Comparison of UMM-Collection Recommend Elements

Highlights

- The Resource Language concept is included in all CMR metadata collections
- The Spatial Representation concept is included in 96% of CMR metadata collections
- The Quality Statement concept is included in 80% of CMR metadata collections

Improvement Focus Areas

- Resource Citation content is missing from the majority of CMR metadata collections
- Resource Access/Use Constraint content is missing from the majority of CMR metadata collections

How are UMM-Collection Optional concepts used in the CMR?

The link below provides a google sheets comparison view of Recommended UMM-Collection concept usage across CMR metadata collections. The pink shaded cells near the bottom of the table indicate recommended elements that are missing for some of the collections.

Table 3: Collection Comparison of UMM-Collection Optional Elements

Highlights

- Sensor Short Name is the most commonly used optional element. It exists in 77% of NASA collections.
- Additional Attributes are the most commonly used optional elements**.** Additional Attributes for describing content Information exist in 44% of NASA collections.
- Additional Attributes also exist in NASA metadata for describing Platform Information, Instrument Information and Quality Information.