GSFC ESDIS CMO
10/11/2017
Released

Revision 1.1
Earth Science Data and Information Systems (ESDIS) Project, Code 423

VARIABLE METADATA MODEL (UMM-VAR)

Goddard Space Flight Center
Greenbelt, Maryland

National Aeronautics and
Space Administration

Signature/Approval Page

Prepared by:

Name Title/Role Organization	Date

Reviewed by:

Name Title/Role Organization	Date

Approved by:

Name Title/Role Organization	Date

Concurred by:

Name Title/Role Organization	Date

[Electronic] Signatures available in B32 Room E148 online at: / https://ops1-cm.ems.eosdis.nasa.gov/cm2/

Preface

This document is under ESDIS Project configuration control. Once this document is approved, ESDIS approved changes are handled in accordance with Class I and Class II change control requirements described in the ESDIS Configuration Management Procedures, and changes to this document shall be made by change bars or by complete revision.

Any questions should be addressed to: esdis-esmo-cmo@lists.nasa.gov

ESDIS Configuration Management Office (CMO)
NASA/GSFC
Code 423
Greenbelt, Md. 20771

Abstract

This document describes the Unified Metadata Model for Variables (UMM-Var) to be used by the NASA Earth Science community. This model takes into account standards and specifications (Directory Interchange Format (DIF), Earth Observing System (EOS) Clearing House (ECHO) 10, ISO 19115-2, and ISO 19115-1) used by this community. Implementers of Earth Science Data and Information System's (ESDIS) CMR, its clients, and data providers should reference this document and the Unified Metadata Model (UMM) as a guide while implementing the system, its clients, or generating metadata.

Change History Log

Revision	Effective Date	Description of Changes (Reference the CCR & CCB Approval Date)
V0.8	July 2015	• Provisional NASA Internal Review Release
V0.9	August 2015	• Provisional NASA Internal Review Release
V1.0	September 2015	• Updated from September 2015 Earth Science Data and Information System (ESDIS) review comments.
V1.1	July 2016 – Sept 2017	• Reworked model to be "Variable" centric versus "Parameter" centric.

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 Purpose

Earth Observing System Data and Information System (EOSDIS) generates, archives, and distributes enormous amounts of Earth Science data via its Distributed Active Archive Centers (DAACs). These data are accessed and employed by a broad user community. It is therefore imperative that reliable, consistent, and high-quality metadata be maintained in order to enable accurate cataloging, discovery, accessibility, and interpretation. To increase the level of quality and consistency among its metadata holdings, EOSDIS has developed a model for various metadata concepts that it archives and maintains. This model aims to document vital elements that may be represented across various metadata formats and standards and unify them through core fields useful for data discovery and service invocations. This unified model, aptly named the Unified Metadata Model, has been developed as part of the EOSDIS Metadata Architecture Studies (MAS I and II) conducted between 2012 and 2013.

The UMM will be used by the CMR and will drive search and retrieval of metadata cataloged within that system.

This document is intended to serve as a reference profile – a part of the UMM model - for geospatial science metadata for collections. This reference profile is referred to as the UMM-Var, where 'Var' indicates that this is the variables profile.

This document provides information to the National Aeronautics and Space Administration (NASA) Earth Science community. Distribution is unlimited.

1.2 Scope

This document is related to the Variable Metadata Model. It describes elements that are specific to this metadata model.

1.3 Impact

This document outlines common elements used in the Unified Metadata Model that are intended to be backward-compatible with existing NASA earth science metadata implementations. It will impact providers from NASA Distributed Active Archive Centers (DAAC[s]), Common Metadata Repository (CMR) client developers, metadata catalog developers, and users.

1.4 Copyright Notice

1.5 Feedback

Questions, comments and recommendations on the contents of this document should be directed to support@earthdata.nasa.gov.

1.6 Document Conventions

This document contains UMM common elements, each of which contains the following components:

- Element Name: Provides the element name.
- Element Specification: Provides the sub-elements, cardinality of the sub-elements within (), any valid values within <>, applicable comments and notes within {}, and any other major factors that make up the element.
- Description: Provides background information on the purpose of the element and its intended use. Furthermore, any information about the element's current usage, recommendations for usage, or unresolved issues is also documented here. The term "resource" is also used mainly in the description section though it can be used elsewhere and it is meant to denote collections, granules, services, visualizations, variables, documents, etc.
- Profile Utilization: Lists which profiles use the specific element.
- Cardinality: Indicates the expectation of counts for this element, summarized in Table 1.
- Analysis: Gives an analysis of this element where needed and describes any necessary reconciliation.
- Mapping: Gives Extensible Markup Language (XML) Path Language (XPath) mappings for this element to the elements in other specifications. This can be considered as the "crosswalk" for this element. For a link to more information about XPaths please see Related Documents.
- Examples: XML snippets from "cross-walked" data standards documenting sample values for the element. Whenever possible, a Uniform Resource Locator (URL) to the specific collection or service used for the metadata snippet, is provided.
- Recommendations: Provides any future recommendations for the element.

Value	Description
1	Exactly one of this element is required
0..N	Optionally, up to and including N numberof this element may be present
0..*	Optionally, any number of this element may be present
1..*	At least one of this element is required, any number may be present

Table 1. Cardinality

1.7 Related Documentation

The latest versions of all documents below should be used. The latest ESDIS Project documents can be obtained from URL: https://ops1-cm.ems.eosdis.nasa.gov. ESDIS documents have a document number starting with either 423 or 505. Other documents are available for reference in the ESDIS project library website at: http://esdisfmp01.gsfc.nasa.gov/esdis_lib/default.php unless indicated otherwise.

1.7.1 Applicable Documents [and Forms]

The following documents are referenced within or are directly applicable, or contain policies or other directive matters that are binding upon the content of this document.

DIF 9	https://gcmd.nasa.gov/Aboutus/xml/dif/dif.xsd https://gcmd.nasa.gov/add/difguide/index.html
DIF 10	https://gcmd.gsfc.nasa.gov/Aboutus/xml/dif/dif_v10.2.xsd https://gcmd.gsfc.nasa.gov/DocumentBuilder/defaultDif10/guide/index.html
ECHO 10	https://wiki.earthdata.nasa.gov/display/echo/Earth+Observing+System+Clearing+House+-+ECHO
SERF	https://gcmd.nasa.gov/Aboutus/xml/serf/serf.xsd https://gcmd.nasa.gov/add/serfguide/index.html
ISO 19115-2 (MENDS)	http://www.iso.org/iso/catalogue_detail.htm?csnumber=39229 https://cdn.earthdata.nasa.gov/iso/
ISO 19115-1	https://github.com/ISO-TC211/XML

Table 2. Applicable Documents

1.7.2 Reference Documents

The following documents are not binding on the content but referenced herein and, amplify or clarify the information presented in this document.

Tags	http://en.wikipedia.org/wiki/Tag_%28metadata%29
XPath	XPath is a language for addressing parts of an XML document, designed for use with XSLT.
XLinks	http://en.wikipedia.org/wiki/XLink

Table 3. Reference Documents

2 UNIFIED METADATA MODEL - VARIABLES

2.1 Introduction to Unified Metadata Model for Variables (UMM-Var)

Earth Observing System Data and Information System (EOSDIS) generates, archives, and distributes massive amounts of Earth Science data via its DAACs. Reliable, consistent and high-quality metadata are essential to enable useful cataloging of these data. To improve the quality and consistency among its metadata holdings, EOSDIS has developed profiles for metadata that it archives and maintains. This model aims to document vital elements that may be represented across various data models and standards and unify them through core fields useful for data discovery and service invocations. This unified model, aptly named the Unified Metadata Model (UMM) will be used by the CMR and will drive search and retrieval of metadata cataloged within that system.

This document describes the Unified Metadata Model for Variables (UMM-Var). It includes the uses cases for UMM-Var, the Variable metadata model itself and its relationship with other UMM models, element descriptions with examples.

Note: Because key terms "Variable" and "Measurement" are frequently used within the Earth Science community, often with different meanings, operative definitions for the purpose of this document are provided below.

Measurement: The act or process of measuring an observable property, usually geophysical. For models, it is a simulated observable property. We describe it by its name in this context.

Examples: Aerosol Optical Depth, Air Temperature, Atmospheric Moisture, Methane, Sulphur Dioxide.

Variable: An artifact that represents a measurement. We describe it by its name and characteristics in this context. The description includes what was being measured (i.e., the observable property, and how it was measured (e.g., what instrument, what resolution, what location, etc.).

Examples: Aerosol Optical Depth 550nm (Dark Target), Aerosol Optical Depth 550nm (Deep Blue, Land Only), Air Temperature (Daytime/Ascending), Air Temperature at 2m, Air Temperature at Surface (Daytime/Ascending), Air Temperature at Surface (Nightime/Descending), Relative Humidity (Daytime/Ascending), Relative Humidity (Nightime/Descending), Water Vapor Mass Mixing Ratio (Daytime/Ascending), Water Vapor Mass Mixing Ratio (Nightime/Descending), Methane Total Column (Nightime/Descending), SO2 Column Mass Density, SO2 Column Mass Concentration.

Sample Illustration

Measurement: Aerosol Optical Depth

Variable: Aerosol Optical Depth 550nm (Dark Target)

Variable: Aerosol Optical Depth 550nm (Deep Blue, Land Only)

Note: The term "Measurement" is the act or process of measuring an observable property, and is mostly likely to be named in any search terms. The term "Variable" is an artifact that represents a Measurement. The description of the Variable includes what was being measured (i.e., the observable property, and how it was measured (e.g., what instrument, what resolution, what location, etc.). The Variable class will be used to store instances of each variable in terms of its metadata. The Variable metadata will consist of its name, and other characteristics. The CMR Variable class can be utilized to simplify search and retrieval of data products at the variable level. Therefore, the CMR will maintain the associations of "Variables" to "Measurements".

In terms of the data product and its file structure, variables are stored within a data granule, e.g. Aerosol Optical Depth 550nm (Dark Target), Aerosol Optical Depth 550nm (Deep Blue, Land Only), along with its associated data quality variables, and ancillary variables, e.g. latitude, longitude information).

2.2 Variable Context Diagram and Metadata Profile Relationships

Figure 1 shows UMM-C metadata class at a high level and its relationships with the other key entities: Collection, Granule, Service, and Variable. It shows specifically a Variable's role in the context of the Unified Metadata Model.

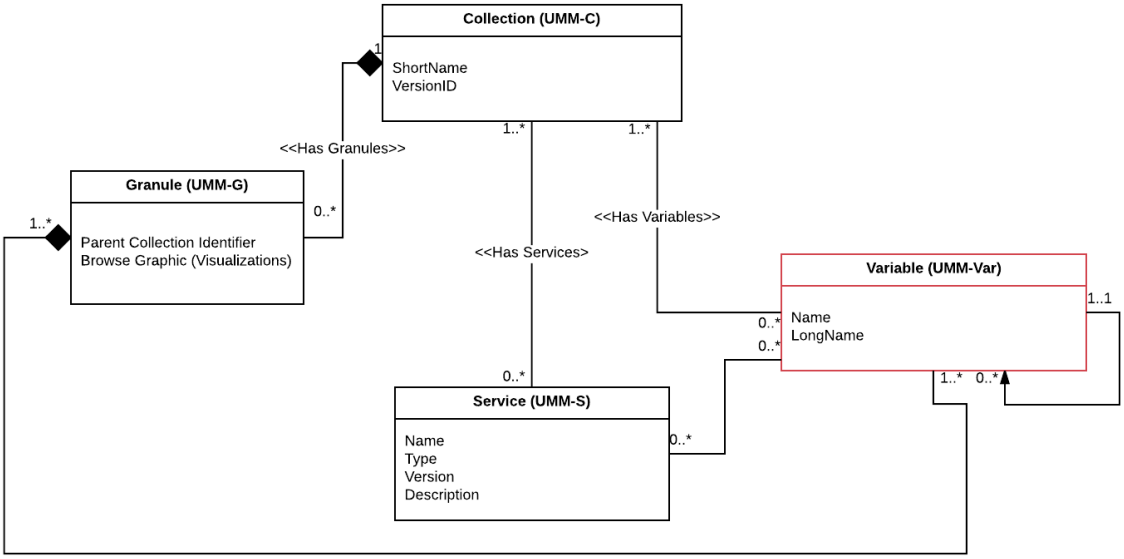


Figure 1. Unified Metadata Model - Variable (UMM-Var) as part of UMM context diagram

2.3 Use Cases

This section provides information about different use cases identified for the UMM-Var. A sequence diagram is presented after each use case.

2.3.1 a. Browse Variables of a Collection

Scenario: The user starts with a collection, and wants to know what variables it includes.

Outcomes: Enables a user without any knowledge of the variable names to search for collections, select one, and be presented with a list of variables for that collection, grouped by measurement.

Use Case: see use case diagram below.

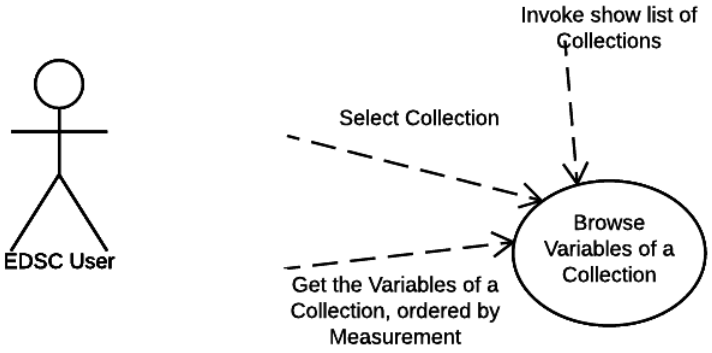


Figure 2. Use Case – Browse Variables of a Collection

Workflow: See sequence diagram below**.**

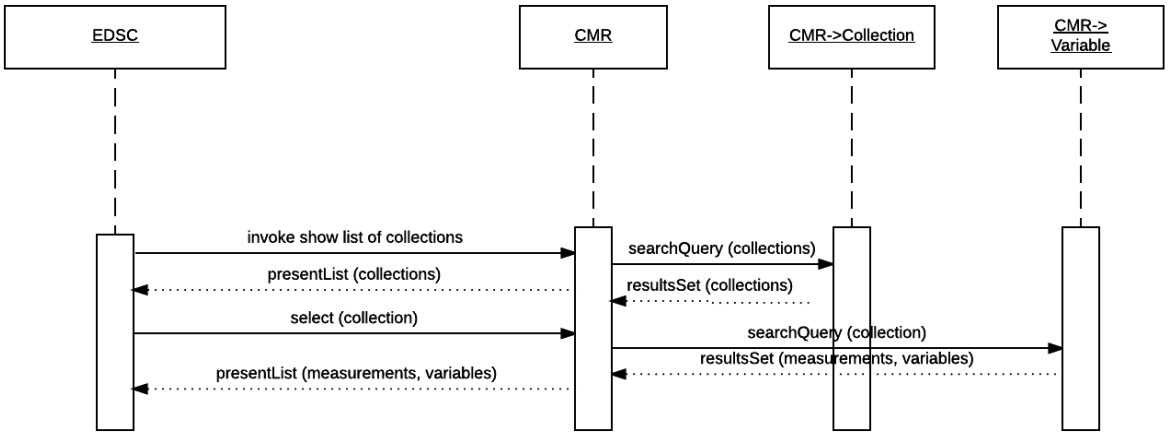


Figure 3. Sequence Diagram – Browse Variables of a Collection

2.3.2 b. Faceted Browse

Scenario [a]: As a user of a search tool such as Earthdata Search Client (EDSC), I can get a list of Measurement facets from the CMR.

Scenario [b]: As a user of the search tool, I can click on a "Measurement" facet value and constrain the lists to the collections that match the selected Measurement and any other constraints I have selected.

Outcomes: The EDSC user, with no knowledge of the Measurements available within the CMR, can get a list of Measurements, and can further constrain the lists to Collections which match by clicking on that Measurement, and any other constraints.

Note: The user can go on and click on a Measurement to see the list of Variables available for that Measurement.

Use Case: see use case diagram below

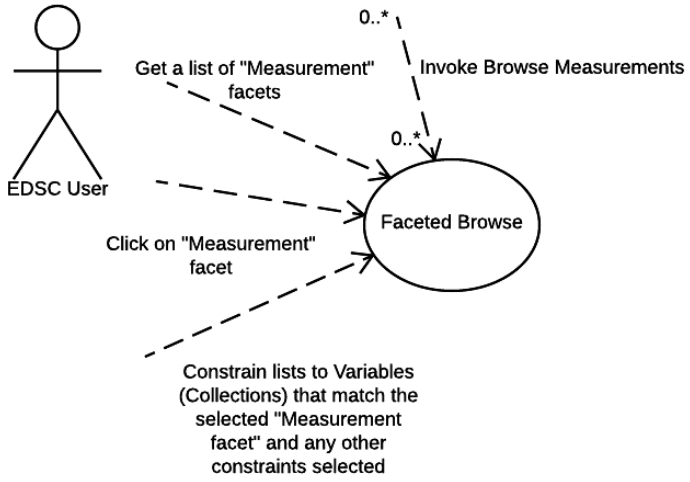


Figure 4. Use Case – Faceted Browse

Workflow: see sequence diagram below

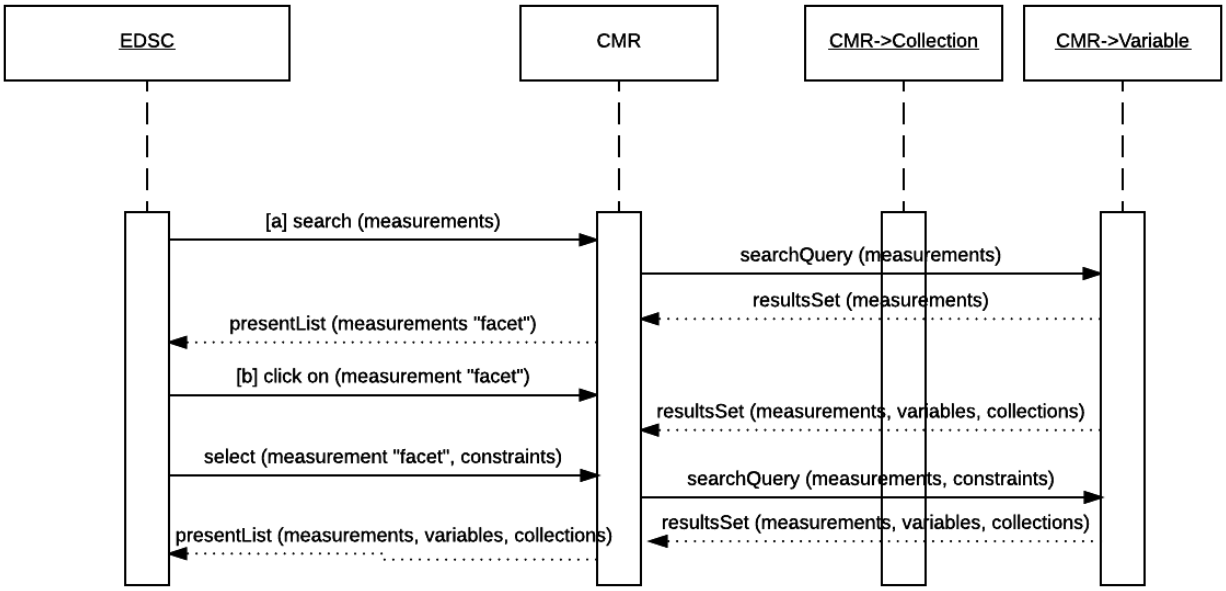


Figure 5. Sequence Diagram – Faceted Browse

2.3.3 c. Update Variable Associations

Scenario [a]: As a CMR client, such as the Metadata Management Tool (MMT), I can associate multiple variables with a collection.

Scenario [b]: As a CMR client, I can submit multiple collections and all of the variables listed for each collection.

Scenario [c]: As a metadata curator, I can populate the list of valid measurements with selections from the GCMD keyword hierarchy, or CSDMS, or CF convention.

Outcomes: The curator seeded the CMR with new valid variable measurements from the GCMD keyword hierarchy and allowed editors to maintain variable and collection associations. A curation tool, e.g. MMT, may be used to maintain/update the variable and collection associations.

Use Case: see use case diagram below.

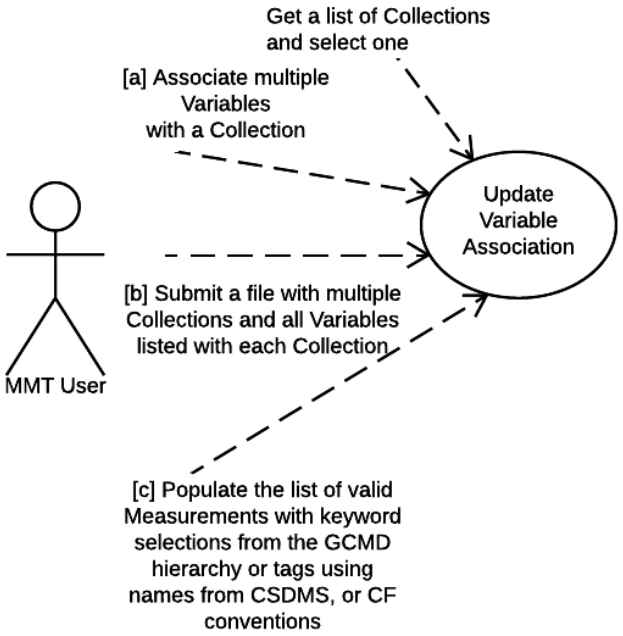


Figure 6. Use Case – Update Variable Associations

Workflow: See sequence diagram below**.**

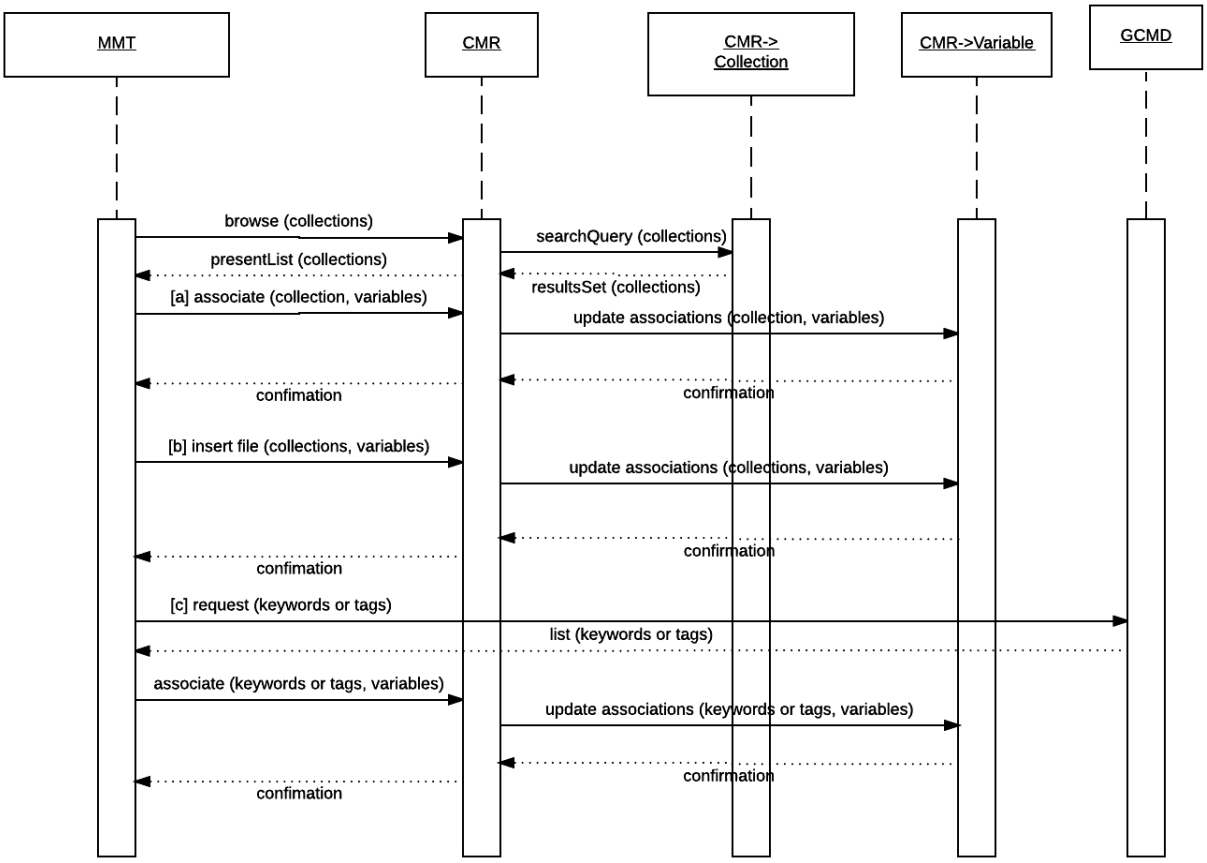


Figure 7. Sequence Diagram – Update Variable Associations

2.3.4 d. Search Relevancy Ranking

Scenario: As a search engine (CMR), I can rank collections with a high relevance ranking when one or more of the search words appear in the measurement names associated with the variables in the collection, as opposed to more generic fields such as the summary or references.

Outcomes: Enables users to search the CMR to get a list of Collections, ranked by relevance to a search word, in matching to a measurement.

Use Case: see use case diagram below**.**

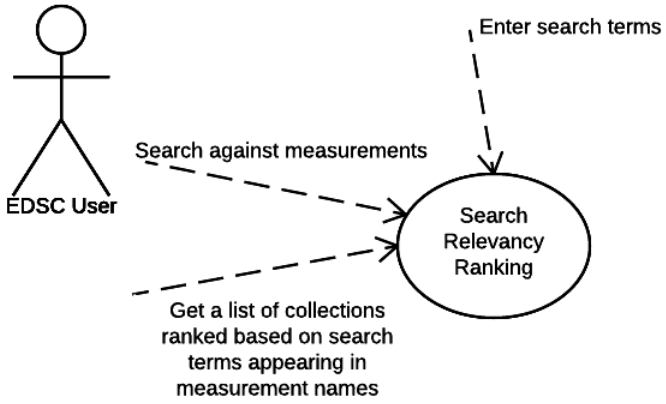


Figure 8. Use Case – Search Relevancy Ranking

Workflow: See sequence diagram below**.**

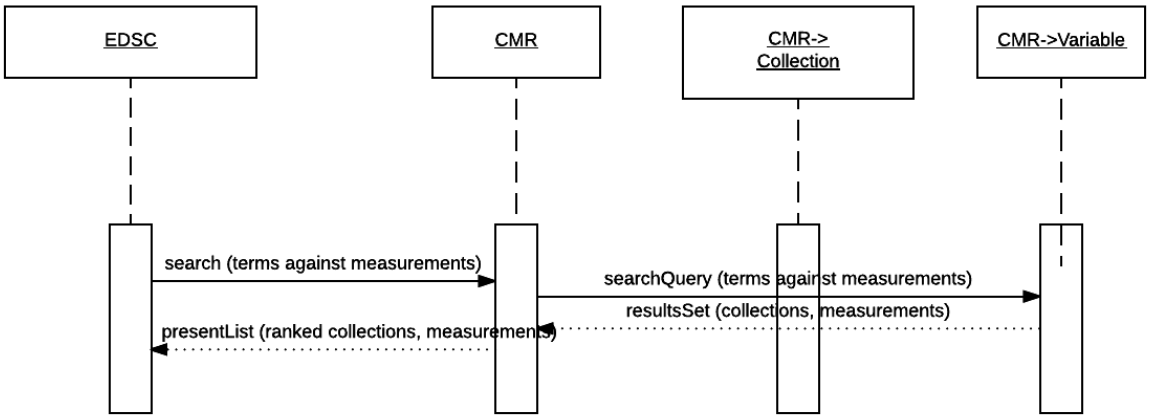


Figure 9. Sequence Diagram– Search Relevancy Ranking

2.3.5 e. Cross-site Data Subsetting

Scenario: As a subsetting GUI, I can present the variables for a given collection in a logically categorized way, such as by measurement, and further subset the data into more specific groups based on additional criteria.

Outcomes: Enables users of a subsetting GUI to perform cross-site subsetting, based on the selection of a collection, categorized by measurement, variables. Cross-site subsetting occurs

when a variable (by its association with a granule) can exist in more than one collection, and these collections may be sourced from multiple sites, (i.e. GES DISC, NSIDC, etc.) The CMR can perform a cross-site search since it houses metadata from all sites. This use case enables a user to go on to perform subsetting, via a GUI.

Note: In the example shown below, the measurement term used was:Ozone. This resulted in two collections being returned from the search: AIRX2RET and OMDOAO3. In the subsetting GUI, variables are shown grouped under the measurement term, for each collection. The user will be able to subset the variable fields, for specific granules of interest.

Use Case: see use case diagram below**.**

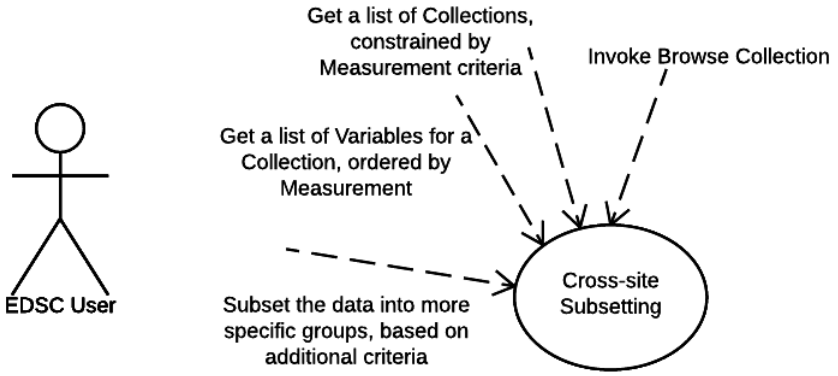
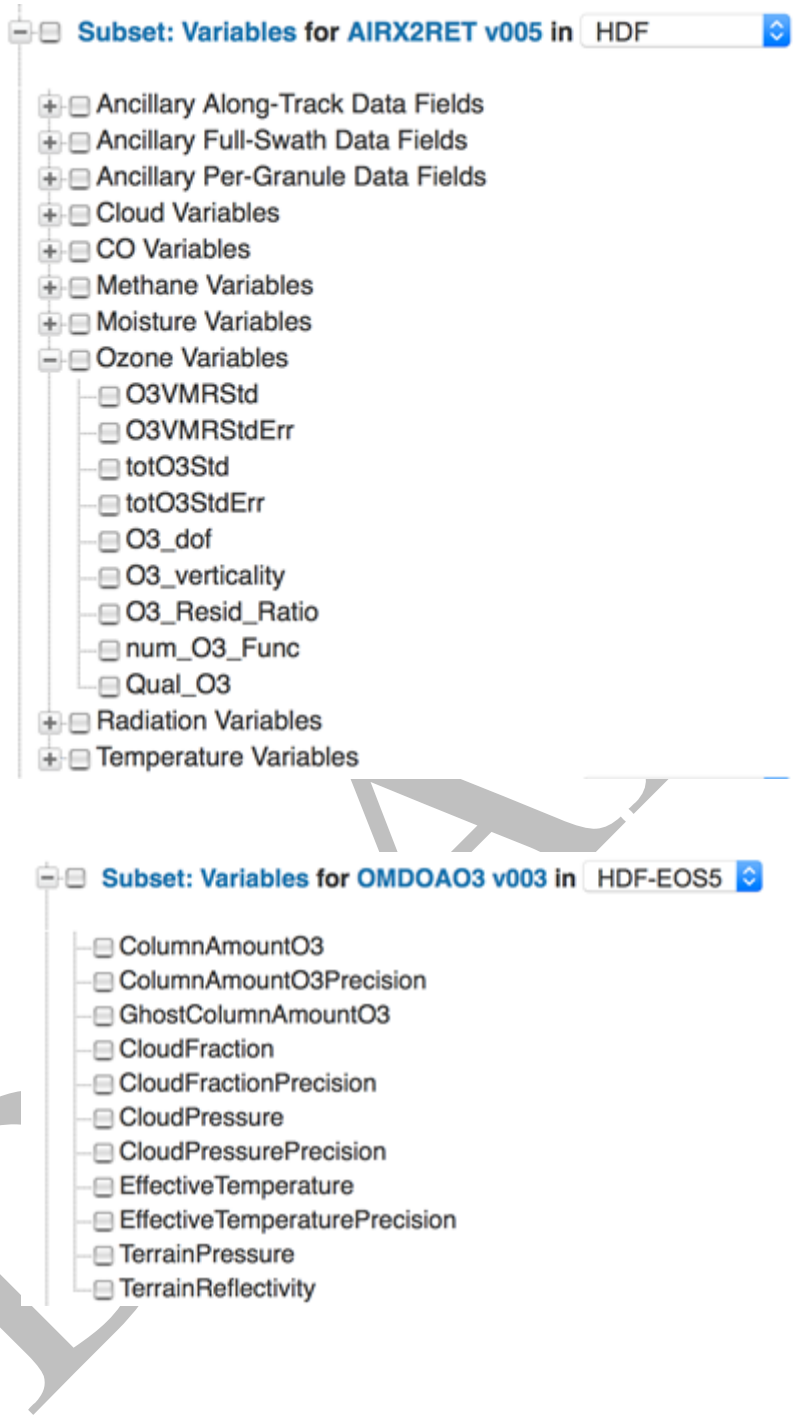


Figure 10. Use Case – Cross-site Data Subsetting



Workflow: See sequence diagram below**.**

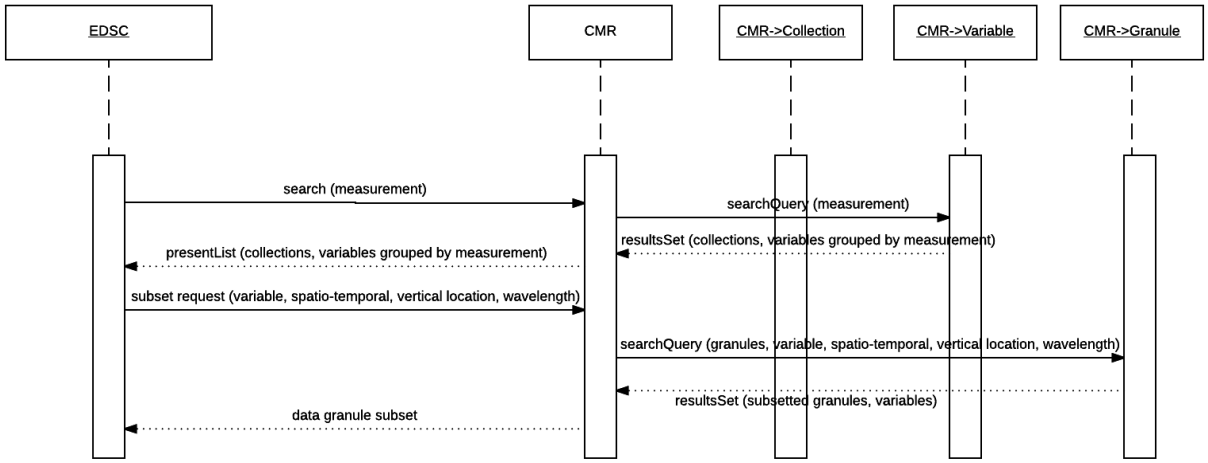


Figure 11. Sequence Diagram – Cross-site Data Subsetting

2.3.6 f. Access Variables Data including Ancillary Variable Data (extension of Cross-Site Data Subsetting Use Case).

Scenario: The user starts with a list of variables (e.g., {<measurement type>, <vertical location>, <wavelength>, ... }, and wants to know which collections contain the specified variables (and may also want to know what data quality, instrument calibration, spacecraft location, etc. variables are needed to properly understand the data).

Outcomes: Enables a user without any knowledge of the collections, to locate those collections which contain a variable that has the properties selected from the initial list of properties. Allows subsequent discovery of associated variables, e.g. ancillary, calibration, location, data quality variables, which are directly related to the science variable selected.

Use Case: see use case diagram below.

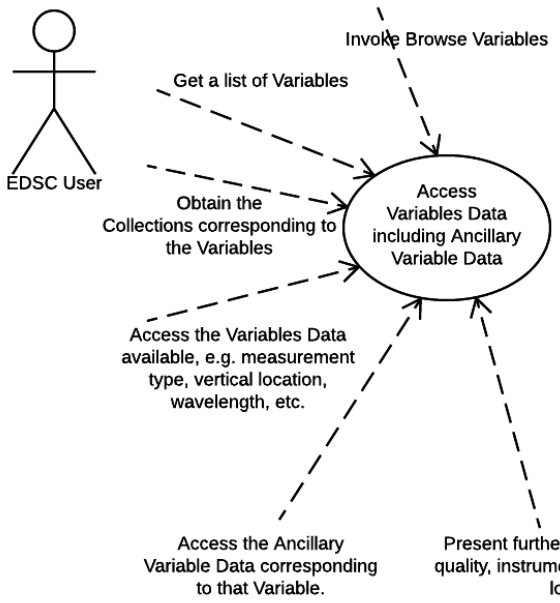


Figure 12. Use Case – Access Variables Data including Ancillary Variable Data

Workflow: See sequence diagram below**.**

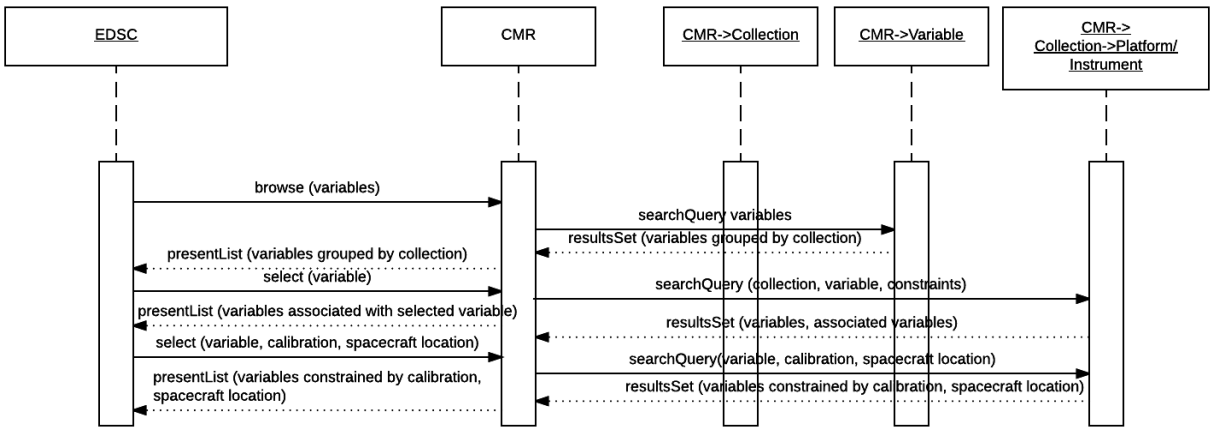


Figure 13. Sequence Diagram – Access Variables Data including Ancillary Variable Data

2.3.7 g. Integrating GIBS with Web-Based Clients

Scenario: As a user of a GIBS client (like EDSC, WorldView, or GloVIS), I can view granules' browse images for a particular layer to form a request to obtain the corresponding data variables, and only those data variables pertaining to this layer. Through CMR, I can locate the granules,

the corresponding data variable from which the layer was generated, and any ancillary variables that need to go along with that variable (coordinates, quality, etc.). Ideally, I can transform that information into a set of subsetting request URLs that will fetch just those data variables from the appropriate granules.

Outcomes: Allow users of a GIBS client to fetch data subsets based on their layer selections, and any associated variables.

Note: It may be a prerequisite for this use case to have a way to invoke "show layers" for a collection, and selected measurement(s), per the GIBS client.

Use Case: see use case diagram below.

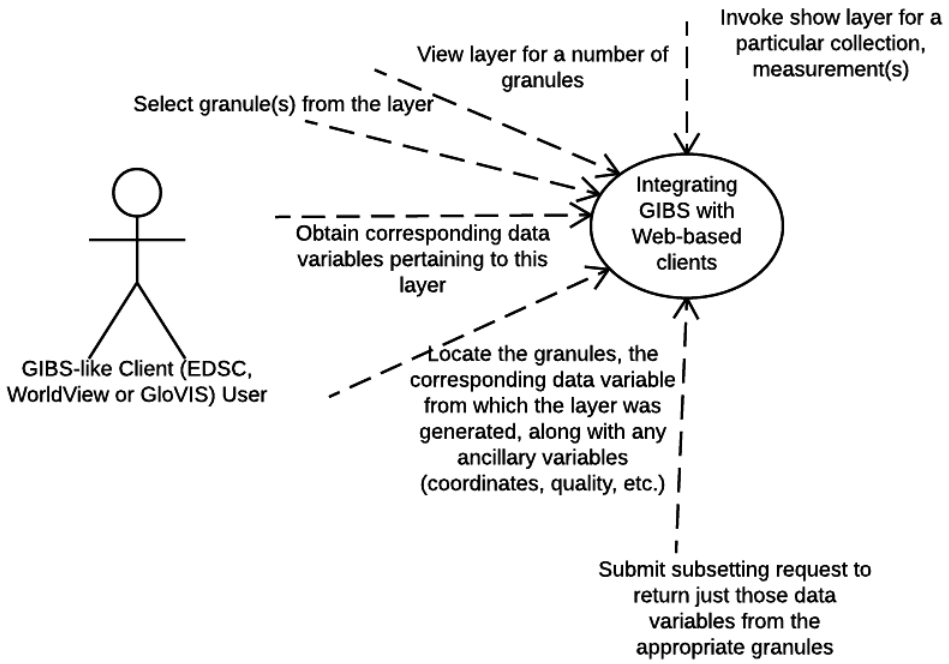


Figure 14. Use Case – Integrating GIBS with web-based clients

Use Case: See use case diagram below**.**

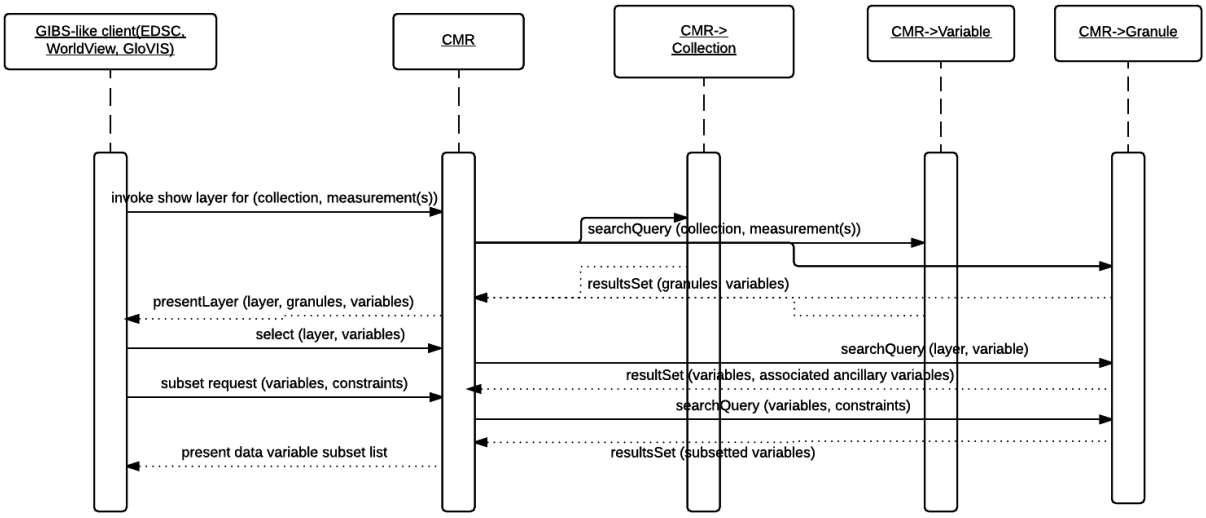


Figure 15. **Sequence Diagram – Integrating GIBS with web-based clients**

2.4 UMM-Var Metadata Model

As shown in Figure 16, the UMM-Var Metadata Model asserts that a Variable metadata instance is related to one or more Collections, one or more Granules, one or more Variables (e.g. a Science Variable may have a related Quality or Ancillary Variable). The remaining classes: Characteristics, ScienceKeywords, Measurements, Set, FillValue, Dimensions and Service are discussed in more detail throughout the remainder of this document. Each relation expresses a different type of information conveyed by the variable.

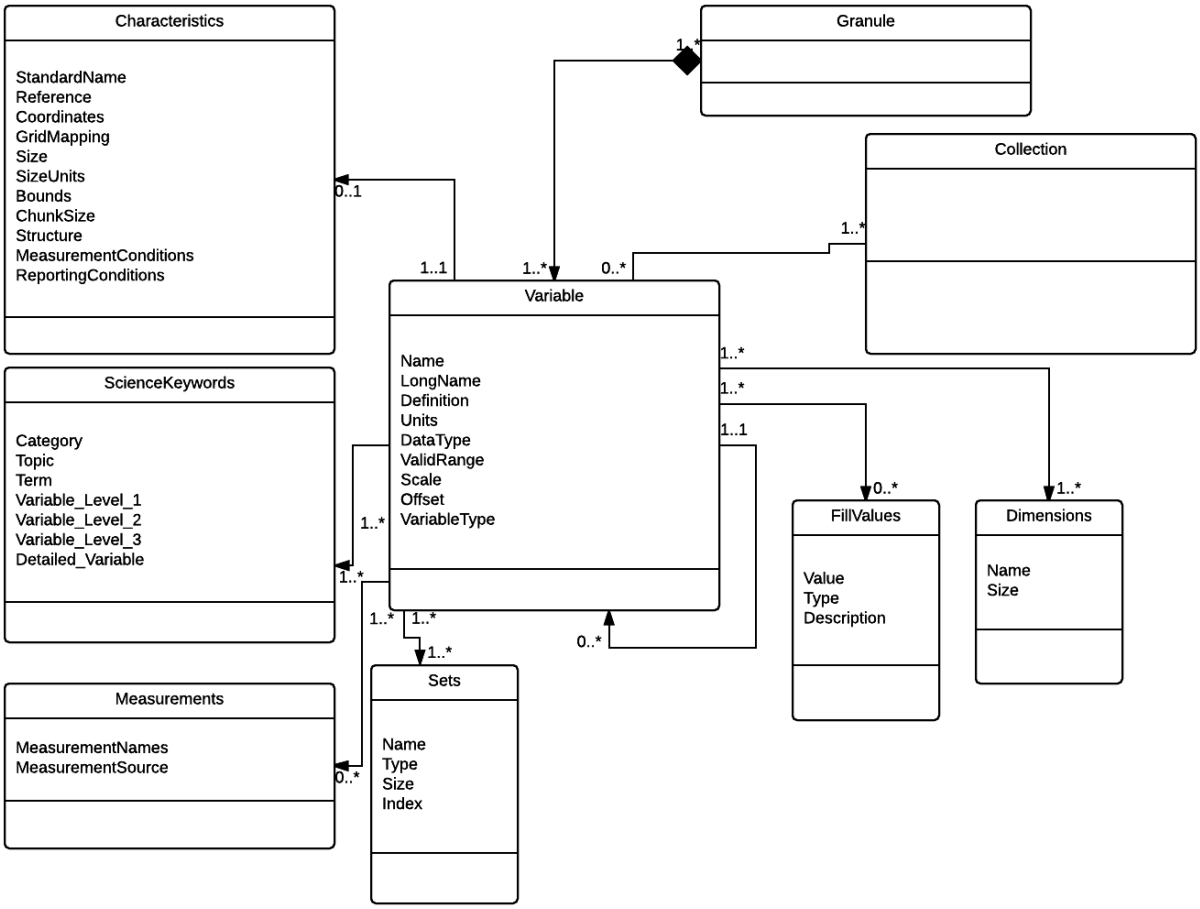


Figure 16. UMM-Var Metadata Model

The author of a Variable metadata record should be cognizant of the following:

1. A Collection has zero or more Variables
2. A Granule aggregates one or more Variables. Note the Variable class lifecycle is dependent on the Granule class instance lifecycle.
3. A Variable may be related to zero or more Variables. For example, a Variable with VariableType: Science may have a related Variable with VariableType: Quality.
4. The elements within the Characteristics section apply to a Variable. Not all Variables have all the elements contained within the Characteristics class.
5. The elements of the ScienceKeywords section also apply to a Variable. The ScienceKeywords may be sourced from GCMD Keywords. See Appendix C.
6. The elements of the Measurement section apply to a Variable. The Measurements may be sourced from the CSDMS standard names or the CF Convention standard names. This process will be dictated by a Governance process. (See Appendix D)
7. The elements of the Service section apply to a Variable. This class is pending definition of the UMM-S model.

8. Information in the Characteristics section should be derived only from the Granule's data file. Granule selected should be from a collection that is associated with the variable.
9. A Variable record may be created / updated via the MMT GUI or XML file.
10. A Variable's record should answer all parts of the following question:

What measurement type, collections, variables, granules are associated with the Variable?

The following element descriptions address these topics.

2.4.1 Variable

Elements

Variable [1..N]
Variable/Name [R]
Variable/LongName [R]
Variable/Definition [R]
Variable/Units [O]
Variable/DataType [R]
Variable/ValidRange [O]
Variable/Scale [R]
Variable/Offset [R]
Variable/VariableType [O]

Note: The required fields: Name, LongName, DataType, Scale, Offset are derived from the Variable fields in the data set. The non-mandatory fields are derived from a candidate sample of the data set, if available. The Definition field is required and this metadata is set by the Metadata Curator via MMT.

2.4.1.1 Name [R]

Element Specification

Variable/Name (1)

Description

Represents the name of a Variable. A Variable is an artifact that represents a measurement. We identify it by its name.This element's value maps to the variable name in the granule's data file.

Sample Value: psi

Tags

Required, Free Text Search

2.4.1.2 LongName [R]

Element Specification

Variable/LongName (1)

Description

The expanded or long name related to the variable Name.

Sample value: mean sea level pressure

Tags

Required, Free Text Search

2.4.1.3 Definition [R]

Element Specification

Variable/Definition (1)

Description

The definition of the variable, in specific terms.

Sample value: mean sea level pressure measured in units of hectopascals.

Tags

Required, Free Text Search

2.4.1.4 Units[O]

Element Specification

Variable/Units (0..1)

Description

The units associated with a variable.

Example: Units = hPa

Tags

Optional

2.4.1.5 DataType [R]

Element Specification

Variable/DataType(1)

Description

Specify basic computer science data type of a variable. These types can be either short, long, character, binary, etc.

Sample value: float , short, byte.

Tags

Required

2.4.1.6 ValidRange [O]

Element Specification

Variable/ValidRange (0..1)
Variable/ValidRange/Max (0..1)
Variable/ValidRange/Min (0..1)

Description

Valid range data value of a variable minimum and maximum value of data value. Due to the inconsistent way collections represent the valid range of a variable, i.e. Min=-300, Max=4500, or valid_range=0.0, 180.0. Some variables carry these as one field with comma delimited values, others separate min and max into two fields. Min and Man values will need to be parsed out.

Sample values: -100, 5000

Tags

Optional

2.4.1.7 Scale [R]

Element Specification

Variable/Scale (1)

Description

Scale factor which has been applied to the variable value, uniformly across the grid or lattice. (ISO 19115-1 definition).

Sample value: 0.00100000004749745

Tags

Required

2.4.1.8 Offset [R]

Element Specification

Variable/Offset (1)

Description

The physical value corresponding to a cell value of zero. (ISO 19115-1 definition).

Sample value: 0.0

Tags

Optional

2.4.1.9 VariableType [O]

Element Specification

Variable/VariableType (0..1)

Description

Specify basic type of a variable. These types can be either:Science, Ancillary, Quality etc.This field is applied by the Metadata Curator, via a suitable GUI, i.e. MMT.

Sample value: Science, Quality.

Tags

Required

2.4.2 Characteristics

The elements of this section apply to a Variable.

Elements

Characteristics (0..1)
Characteristics/StandardName[O]
Characteristics/Reference[O]
Characteristics/Coordinates [O]
Characteristics/GridMapping [O]
Characteristics/Size [O]
Characteristics/SizeUnits [O]
Characteristics/Bounds[O]
Characteristics/ChunkSize [O]
Characteristics/Structure [O]
Characteristics/MeasurementConditions [O]
Characteristics/ReportingConditions [O]

Note: StandardName is derived from the Variable fields in the data set, if available. Most of the elements of this class are not always available, but typical across collections.

2.4.2.1 StandardName[O]

Element Specification

Characteristics/StandardName (0..1)

Description

The CF-compliant "Standard Name" for the variable. See CF Standard Names.

Sample value: atmosphere_absorption_optical_thickness_due_to_ambient_aerosol_particles

Tags

Optional, Free Text Search

2.4.2.2 Reference[O]

Element Specification

Characteristics/Reference(0..1)

Description

Provides a link to variable's reference documentation.

Sample value: http://disc.sci.gsfc.nasa.gov/giovanni/additional/users-manual/G3_manual_parameter_appendix.shtml#deepblue

Tags

Optional

2.4.2.3 Coordinates [O]

Element Specification

Characteristics/Coordinates (0..1)

Description

A text description of the variable’s coordinate range as given by the data provider. For example, '90N, 90S, 180E, 180W'.

Sample value: lonlat

Tags

Optional

2.4.2.4 GridMapping[O]

Element Specification

Characteristics/GridMapping(0..1)

Description

A text description of the variable's mapping projection standard for the variable. For example: 'WGS84 Web Mercator'.

Sample value: WGS84 Web Mercator

Tags

Optional

2.4.2.5 Size [O]

Element Specification

Characteristics/Size {0..1)

Description

The computed byte size for the variable, per the data field. Typically, this is the X dimension x the Y dimension x the number of bytes in the data type (8, 16, 32, etc.)

Sample values: e.g. 2000 x 1000 x 16 = 32,000,000. This field is typically derived from the data set.

Tags

Optional

2.4.2.6 SizeUnits[O]

Element Specification

Characteristics/SizeUnits (0..1)

Description

Computed byte size for the variable, per the data field.

Sample values:

B, KB, MB, GB

Tags

Optional

2.4.2.7 Bounds [O]

Element Specification

Characteristics/Bounds (0..1)

Description

Describe the spatial bounds of a variable. Western Longitude, Northern Latitude and Eastern Longitude, Southern Latitude - specified in decimal degrees.

Sample value: UpperLeftPointMtrs = -180.0, 89.5; LowerRightMtrs = 177.5, -89.5;

Tags

Optional

2.4.2.8 ChunkSize [O]

Element Specification

Characteristics/ChunkSize (0..1)

Description

Describe the chunk size of a variable.

Sample value: 100

Tags

Optional

2.4.2.9 Structure [O]

Element Specification

Characteristics/Structure (0..1)

Description

The full path to the variable within the Granule. For example, '/<filename>/MODIS_Grid_Daily_1km_LST/Data_Fields/', where filename = 'MOD11A1.A2009172.h16v05.006.2016014073638.hdf'.

Tags

Optional

2.4.2.10 MeasurementCondition [O]

Element Specification

Characteristics/MeasurementCondition (0..1)

Description

The measurement conditions of the variable. For example, 'Sampled Particle Size Range: 90 - 600 nm'.

2.4.2.11 ReportingCondition [O]

Element Specification

Characteristics/ResportingCondition (0..1)

Description

The reporting conditions of the variable. The conditions over which the measurement of the variable are valid. For example, 'STP: 1013 mb and 273 K'.

2.4.3 Measurement Information

Elements in this category are used for search purposes.

Elements

Measurements [0.N]
Measurements/MeasurementNames[O]
Measurements/MeasurementSource[O]

Measurements are added to a Variable by a Metadata Curator via a suitable GUI, or manually in the metadata file. Measurements are derived from an uncontrolled source, e.g. CSDMS Standard Names, or CF Convention Standard names, or other sources. Searchable names are provided in this class to enable better search by human-readable Measurement terms, which can be associated with each Variable.

2.4.3.1 MeasurementNames [O]

Element Specification

Measurements/MeasurementNames (0..N)

Description

The names of the measurement derived from science disciplines associated with this variable. The Metadata Curator will apply the Measurements via a suitable GUI, i.e. MMT, or manually in the metadata file.In the future, these will come from either Community Surface Dynamics Modeling System (CSDMS) or Climate and Forecast (CF) Standard Name Table. See Appendix C.

Sample Values:

Aerosols, Land, Mean, Corrected, Air Temperature.

Tags

Required, Uncontrolled Vocabulary

2.4.3.2 MeasurementSource [O]

Element Specification

Measurements/MeasurementSources (0..N)

Description

The specific source of the measurements. Measurements will come from either Community Surface Dynamics Modeling System (CSDMS) or Climate and Forecast (CF) Standard Name Table, or other sources, e.g. Atmospheric Sciences community. See Appendix C.

Sample Values:

CSDMS, CF, BODC.

Tags

Required, Uncontrolled Vocabulary

2.4.4 ScienceKeywords

Elements in this category are used for search and faceting purposes.

Elements

ScienceKeywords (1..N)
ScienceKeywords/Category [R]
ScienceKeywords/Topic [R]
ScienceKeywords//Term [R]
ScienceKeywords/Variable_Level1 [O]
ScienceKeywords/Variable_Level2 [O]
ScienceKeywords/Variable_Level3 [O]
ScienceKeywords/Variable_Level4 [O]
ScienceKeywords/Detailed_Variable [O]

Science Keywords are added to a Variable by a Metadata Curator via a suitable GUI, e.g. MMT, or manually in the metadata file. Science Keywords are derived from a controlled source, e.g. GCMD Keywords. Science Keywords are provided to enable better search by human-readable Measurement terms, which can be associated with each Variable. Note that GCMD Keywords are hierarchical and have a more complex structure than the alternate Measurement class.

Note: Science Keywords search is offered as the primary way to achieve a Measurement search. ScienceKeywords and Measurements could both be used for faceted browse in search clients.

2.4.4.1 ScienceKeywords [R]

Element Specification

ScienceKeywords (1..N)

ScienceKeywords/Category [R]
ScienceKeywords/Topic [R]
ScienceKeywords//Term [R]
ScienceKeywords/Variable_Level1 [O]
ScienceKeywords/Variable_Level2 [O]
ScienceKeywords/Variable_Level3 [O]
ScienceKeywords/Variable_Level4 [O]
ScienceKeywords/Detailed_Variable [O]

Description

Controlled Science Keywords describing the measurements/variables. The controlled vocabulary for Science Keywords is maintained in the Keyword Management System (KMS). These will be sourced from GCMD Keywords. See Appendix C.

Sample Values:

"Category": "EARTH SCIENCE", "Topic": "ATMOSPHERE", "Term": "ATMOSPHERIC CHEMISTRY", "VariableLevel1": "NITROGEN COMPOUNDS", "VariableLevel2": "Peroxyacyl Nitrate".

Tags

Required, Controlled Vocabulary

2.4.5 Sets Information

Elements in this category are used to group variables into sets.

Elements

Sets [1..N]
Sets/Name [R]
Set/sType [R]
Set/Size [R]
Sets/Index [R]

The set information of a variable. The variable is grouped within a set. The set is defined by the name, type, size and index. For example, Name: 'Data_Fields', Type: 'General', Size: '15', Index: '7' for the case of the variable named 'LST_Day_1km'.

2.4.5.1 Name [R]

This element enables specification of set name. For example, 'Data_Fields'.

2.4.5.2 Type [R]

This element enables specification of set type. For example, if the variables have been grouped together based on a particular theme, such as wavelength, then the type should be set to that theme, otherwise it should be set to 'General'.

2.4.5.3 Size [R]

This element specifies the number of variables in the set. For example, if the number of variables in the set is fifteen, the size should be set to '15'.

2.4.5.4 Index [R]

This element specifies the index value within the set for this variable, For example, if this variable is the third variable in the set, the index value should be set to '3'.

2.4.6 FillValues Information

FillValues [0..N]
FillValues/Value [R]
FillValues/Type [R]
FillValues/Description [O]

The fill value of the variable in the data file. It is generally a value which falls outside the valid range. For example, if the valid range is '0, 360', the fill value may be '-1'. The fill value type is data provider-defined. For example, 'Out of Valid Range'.

2.4.6.1 Value [R]

Element Specification

FillValue/Value (1)

Description

The fill value of the variable in the data file.

Sample values:

-9999.

Tags

Required

2.4.6.2 Type [R]

Element Specification

FillValue/Type (1)

Description

Type of the fill value of the variable in the data file.

Sample values:

Science

Tags

Required

2.4.6.3 Description [O]

Element Specification

FillValue/Description (0..1)

Description

Description of the fill value of the variable in the data file.

Sample values: Valid Science Fill Value

Tags

Optional

2.4.7 Dimensions Information

Dimensions [1..N]
Dimensions/Name [R]
Dimensions/Size[R]

A variable consists of one or more dimensions. An example of a dimension name is 'XDim'. An example of a dimension size is '1200'. Variables are rarely one dimensional.

2.4.7.1 Name [R]

Element Specification

Dimensions/Name(1..N)

Description

The name of the dimensions of the variable represented in the data field. For example, 'XDim'.

Sample values:

SurfacePressure, XDim, YDim

Tags

Required

2.4.7.2 Size [R]

Element Specification

Dimensions/Size(1..N)

Description

The size of the dimensions of the variable represented in the data field. For example, '1200'.

Sample values:

1200, 3600

Tags

Required

Appendix A: Tags Glossary

The following table lists all tags used in this model and provides a description of the tags' usage.

Tag Name	Description
[O]ptional	This field is not required.
[R]equired	This field is required.
[C]onditional	This field is required if some conditions are met.
Controlled Vocabulary	This field is populatedfrom a vocabulary that is used to validate the value. The validation task will most likely be performed via a vocabulary management service.
Free Text Search	This field is indexed by the CMR as part of the free text search.
Search API	This field will be indexed by the CMR. The CMRwill be exposed via a search API.
Deprecate	This field should be deprecated in a future revision of the UMM-Var.
Faceted	This field should be exposed by the CMR catalog via a faceted search response.
Markdown Support	This class supports markdown-formatted text. Additional information on markdown can be found at http://en.wikipedia.org/wiki/Markdown
Validated Relationship	Any associated granules use this field for validation against the collection, or associated collections are checked for validity

Table 4. A list of all tags with descriptions

Appendix B: Inputs

- Discussions with Kathleen Baynes, Jon Pals (ECS Science Office), Dan Pilone (EED Chief Technologist), Abe Taaheri (HDF-EOS Lead Developer), Christopher Lynnes (NASA/ESDIS), Mahabal Hegde (NASA/GSFC)
- Inputs from Kathleen Baynes
- Original App-friendly EOSDIS Science Information Retriever (AESIR) proposal: https://docs.google.com/document/d/1H0EAPV3Iv3WtheOAVZrJ1eY89NneyiweCcHAz5eI9a0/edit#heading=h.8r4j53o9wk3j
- Variable data use case: http://giovanni.gsfc.nasa.gov/giovanni/
- Variable response from AESIR: AESIRresponse.xml file
- ODISEES tool link: http://odisees.larc.nasa.gov/
- (Point of Contact for this is Beth Huffer: https://wiki.earthdata.nasa.gov/display/~bhuffer)
- More information about the GIBS API: https://wiki.earthdata.nasa.gov/display/GIBS/GIBS+API+for+Developers
- Existing/Preliminary UMM-V Model: https://docs.google.com/spreadsheets/d/1Zm5i_Mln0jGGHDS9wC8vzNZAqZ034bu5aSk58Bn77Ig/edit#gid=0 (to be refined by the GIBS Technical Working Group)
- GIBS TWG Wiki Space: https://wiki.earthdata.nasa.gov/display/GIBS/GIBS+Technical+Working+Group
- Current Status of UMM-V: https://wiki.earthdata.nasa.gov/display/GIBS/Visualization+Metadata

- Inputs from Edward Seiler regarding AESIR questions and content
- Giovanni application : http://giovanni.gsfc.nasa.gov/giovanni/
- The information for each variable is populated via an editor named EDDA at
- http://dev-ts1.gesdisc.eosdis.nasa.gov/EDDA/index.html
- The help page for EDDA at
- http://dev-ts1.gesdisc.eosdis.nasa.gov/EDDA/EDDA_help.html
- SMAP Data – sample 'Freeze Thaw' variable data from Abe Taaheri: SMAP_L3_FT_A_20140122_R11160_001.h5
- CF Conventions and CF Standard Variable Names: http://cfconventions.org/Data/cf-standard-names/27/build/cf-standard-name-table.html
- CSDMS Standard Variable Names: http://csdms.colordo.edu/wiki/CSN_Examples
- The CSDMS Standard Names: Cross-Domain Naming Conventions for Describing Process Models, Data Sets Their Associated Variables, S.D. Peckham, University of Colorado, INSTAAR, 1560 30th Street, Boulder, CO (Scott.Peckham@colorado.edu) http://www.iemss.org/sites/iemss2014/papers/iemss2014_submission_263.pdf
- CMR Data Partner User Guide: https://wiki.earthdata.nasa.gov/display/CMR/CMR+Data+Partner+User+Guide#CMRDataPartn
- Towards Unifying NASA Earth Science Enterprise-Wide Metadata Around International Standards: Study Results and Recommendations, S.J.S. Khalsa, CIRES, University of Colorado, Boulder, CO 80309 USA – sjsk@nsidc.org, S.F. Browdy, OMS Tech, Orlando, Florida, USA – steveb@omstech.com, B.H. Wiess, Jet Propulsion Laboratory, Pasadena, CA – Barry.h.weiss@jpl.nasa.gov http://www.isprs.org/proceedings/2011/ISRSE-34/211104015Final00852.pdf

Appendix C: Keywords and Measurements Governance Structure

The Governance Structure shown in Figure 17 for the selection of science keywords and measurements. ESDIS chairs each of the measurement or science keyword selection councils and provides overall science guidance, and the DAAC/Data Providers serve as the decision authority for the metadata associated with data sets sourced from their DAAC/Project.

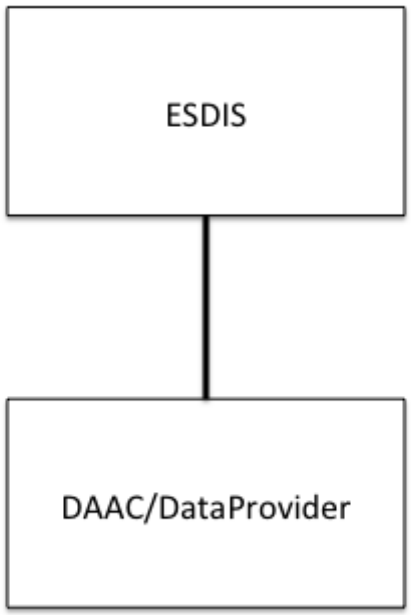


Figure 17. Governance Structure

Adding Science Keywords or Measurements are to be done by the Metadata Curator, via a GUI. Science Keywords are to be sourced from the GCMD Keywords and are controlled. What is being proposed here is not too different from the existing method used in the EDSC UI, with the exception being that the Science Keyword will be used for discovery at the Variable level, as opposed to the Collection or Granule level, which is currently the case. The challenge with Measurements is that they are uncontrolled. The concept is to start with a pre-seeded list of suggested measurements and, over time an alphabetically ordered list can be collected, by certain users, and by their use of a metadata management tool, e.g. MMT. The guidelines for adding science keywords or measurements can be achieved by following the suggested steps below.

Science Keywords

Science Keywords may be selected from the GCMD Keywords set.

The GCMD Keywords are already subject to a strict governance process:

1. Review the controlled keyword/guidelines located at: http://gcmd.nasa.gov/learn/rules.html
2. Verify that the keyword does not already exist.
3. Map these to the appropriate variables.
4. Include a definition of the controlled keyword.

Measurements

Measurements may be selected from an array of standard sources, e.g. CSDMS, CF Conventions, etc.

The process by which measurements may be selected is simple.

1. Determine level: i.e. Atmosphere, Oceans, Land (highest) or Atmosphere Air Temperature (mid), or Atmosphere Air Temperature Saturated Adiabatic Lapse Rate (lowest), etc.

2. Determine whether the measurement is missing, and a new one is needed.

For example, if we have Atmosphere Air Column Water Vapor and the next tag is Atmosphere Air Flow Azimuth Angle of Bolus Velocity, then Atmosphere Air Carbon Dioxide (and its derivatives) are missing.

3. Select the most appropriate measurement to suit the need. If the measurement does not exist, apply crosswalk to another standard, i.e. CSDMS to CF convention Standard Names.

4. Add to the measurements list stored in the CMR so that all future users can use this measurement.

5. Map these to the appropriate variables.

6. Include a definition of the uncontrolled measurement.

Appendix D: Analysis of CSDMS and CF Standard Names as a Source of Tagging

Analysis of CSDMS Standard Names as a source of tagging

The CSDMS (Community Surface Dynamics Modeling System) modeling framework provides mechanisms that allow models and data sets from different contributors (i.e. from different geoscience domains: hydrology, oceanography, meteorology, seismology). The framework defines an approach to the semantic mediation problem. It offers a unique approach to solving this problem by offering a set of standardized and precise descriptions of each variable. It provides a holistic approach to solve the semantic mediation problem by giving a number of options to resolving which names and abbreviations are to be used for a variable.

The naming conventions of the CSDMS Standard Names are based on object-oriented principles.

CSDMS Standard Names are grouped, i.e. by Variables Names for: the Atmosphere, Atoms, Automobiles, Basins, Bedrock, Channel, Chocolate, Compounds and Mixtures, Earthquakes, Glaciers, Materials, Models, Molecules, Oceans, Planets, Projectiles, Radiation, River Deltas, Sea Ice, Snow, Soil, Sea Floor Debris, Topography and Water Tank.

Only a subset of these are groups are suitable for EOS: Variables Names for: the Atmosphere, Oceans, Radiation, Sea Ice, Soil, Snow, Topography.

Potentially, other groups are suitable for EOS: Variable Names for: Basins, Channel, Earthquakes, Glaciers, Planets, River Deltas, Sea Floor Debris.

These Standard Names can be chosen as the primary source of tagging, since each group is highly relevant to the science domains which are covered by the EOS data sets and those likely to be covered in the future.

The CSDMS Standard Names exhibit the Object Name + Model Name Pattern structure to the name.

An example of an Object Name is:atmosphere_water

Examples of the corresponding Model Name Patterns are: domain time integral of precipitation leq volume flux, icefall mass per volume density, precipitation duration, precipitation leq volume flux, precipitation mass flux.

The first example of combining Object Name + Model Name yields the resultant standard name: atmosphere water domain time integral of precipitation leq volume flux.

An example of a missing name would be: atmosphere water precipitation, or the more common term, precipitation.

Analysis of CF Standard Names as a source of tagging

CF conventions for climate and forecast metadata are designed to promote the processing and sharing of files created with the netCDF Application Programmer Interface. The CF conventions generalize and extend the COARDS conventions.

Most of the CF standard names have been derived from guidelines which have drawn on ECMWF, and NCEP GRIB tables, the PCMDI and GCMD.

CF standard names consist of lower-letters, digits and underscores, and begin with a letter. Upper case is not used.

US spelling is used, e.g. vapor, sulfur.

The CF Standard Names can be chosen as a source of tagging supplementary to CSDMS.

Examples of CF Standard Names are: precipitation amount, and precipitation flux, and precipitation flux onto canopy are included in the CF Standard Names and not in the CSDMS Standard Names.

In this simple example, both CSDMS and CF Standard Names may be used as a source of tagging for search terms to locate all variables associated with the measurement: precipitation.

Appendix E: Abbreviations and Acronym List

ACL - Access Control List
AESIR - App-friendly EOSDIS Science Information Retriever
CF - Climate and Forecast metadata
CMR - Common Metadata Repository
COARDS - Conventions for the standardization of NetCDF files
DAAC - Distributed Active Archive Center
DOI - Digital Object Identifier
ECHO - Earth Observing System (EOS) Clearing House
ECMWF - The European Center for Medium-Range Weather Forecasts
ECS - EOSDIS Core System
EDSC - Earthdata Search Client
EED - EOSDIS Evolution and Development
EOS - Earth Observing System
EOSDIS - Earth Observing System Data and Information System
ESDIS - Earth Science Data and Information System
ESI - EOSDIS Service Interface
GCMD - Global Change Master Directory
GRIB - GRIdded Binary file format
ISO - International Organization for Standardization
KMS - Keyword Management System

MAS - Metadata Architecture Studies
MENDS - Metadata Evolution for NASA Data Systems
NASA - National Aeronautics and Space Administration
NCEP - National Center for Earth
NOAA - National Oceanic and Atmospheric Administration
PCMDI - Program for Climate Model Diagnosis and Intercomparison
SERF - Service Entry Resource Format
SSW - Simple Subset Wizard
TBS - To Be Supplied
UML - Unified Modeling Language
UMM - Unified Metadata Model
UMM-C - Unified Metadata Model - Collections
UMM-D - Unified Metadata Model - Documents
UMM-G - Unified Metadata Model - Granules
UMM-S - Unified Metadata Model - Services
UMM-Var - Unified Metadata Model - Variables
UMM-Vis - Unified Metadata Model - Visualization
URI - Uniform Resource Identifier
URL - Uniform Resource Locator
XML - Extensible Markup Language
XPath - XML Path Language