MET Norway Metadata Format Specification

Øystein Godøy, Trygve Halsne, Øystein Torget, Bard Saadatnejad, Egil Støren, Sarfraz Alam

Version 2.2 (draft)

Table of Contents

Abstract	1
Revision history	2
1. Introduction	4
1.1. Applicable documents	4
1.2. Scope	4
1.3. Objectives	5
2. Metadata elements	6
2.1. metadata_identifier	6
2.2. last_metadata_update	6
2.3. metadata_status	7
2.4. collection	7
2.5. title	8
2.6. abstract	9
2.7. temporal_extent	
2.8. geographic_extent/rectangle	12
2.9. geographic_extent/polygon	12
2.10. location	13
2.11. dataset_production_status	15
2.12. dataset_language	15
2.13. operational_status	
2.14. access_constraint	
2.15. use_constraint	17
2.16. personnel	17
2.17. data_center	19
2.18. data_access.	
2.19. related_dataset	21
2.20. related_information.	22
2.21. iso_topic_category	23
2.22. keywords	24
2.23. project	25
2.24. platform	
2.25. instrument	27
2.26. activity_type	28
2.27. reference	28
2.28. dataset_citation	
2.29. cloud_cover	32
2.30. scene_cover	33
3. Explanations	

	3.1. Interpretation of child and parent relationship	35
4.	Controlled vocabularies	36
	4.1. Collection keywords	36
	4.2. Dataset production status	36
	4.3. ISO to MMD mapping.	37
	4.4. MMD to ISO mapping.	37
	4.5. Operational status	37
	4.6. Access constraints.	38
	4.7. Use constraints	38
	4.8. Activity type	38
	4.9. Variable/parameter descriptions	40
	4.10. Platform	40
	4.11. Instruments	41
	4.12. Instrument modes	41
	4.13. Polarisation modes.	41
	4.14. Product type.	42
	4.15. Contact roles	42
	4.16. MMD to ISO mapping	42
	4.17. ISO Topic categories.	43
	4.18. Related Information types	44
	4.19. Data Access Types.	45

Abstract

Revision history

Version	Date	Comment	Responsible
2.2	2019-10-??	Several minor modifications has not been document. Collecting these and preparing for transition into ASCIDOC like CF-conventions.	Øystein Godøy
2.1	2018-12-13	Modifications of controlled vocabularies related to satellite data and data collections, minor edit on related_dataset, major edit on reference.	Øystein Godøy
2.0	2017-10-17	Added controlled vocabularies and elements for instrument children like mode, polarisation and product_type.	Trygve Halsne Øystein Godøy
1.9	2017-05-29	Added ODATA to data_access to support NBS.	Øystein Godøy Egil Støren Bard Saadatnejad Trygve Halsne
1.8	2017-03-13	Adaptations to support NBS (in particular Sentinel-2). Depreciation of system_specific elements used for Halo and BarentsWatch (covered by other elements).	Øystein Godøy Sarfraz Alam Trygve Halsne Bard Saadatnejad
1.7	2016-11-04	Consolidation of document, addition of collection keyword and preparation for NBS usage.	Øystein Godøy Sarfraz Alam

Version	Date	Comment	Responsible
1.6	2016-05-10	Minor formatting changes. Addition of configuration metadata elements (from XMD and others). Addition of vocabulary suggestions.	Øystein Godøy Bard Saadatnejad Sarfraz Alam
1.5	2013-07-15		Øystein Torget
1.4	2013-06-20		Øystein Torget
1.3	2013-03-13		Øystein Torget
1.2	2013-03-06	Added Obsolete as a dataset status. Re-formatted document and added some more formalism to the intro sections.	Øystein Godøy, Øystein Torget
1.1	2013-02-12	Added mapping between MMD and ISO dataset status codes. Move vocabulary from sub element of keywords to attribute.	Øystein Torget
1.0	2013-01-18	First version.	Øystein Torget

Chapter 1. Introduction

Metno Metadata Format (MMD) is an XML metadata format for storing information about scientific datasets. It is meant for consumption by internal systems and to be a corner stone in our data management policy.

1.1. Applicable documents

- 1. ISO19115
- 2. ISO19115 North American Profile
- 3. Environmental Climate Data Sweden ISO19115 Profile
- 4. INSPIRE Metadata
- 5. WMO Core Profile (ISO19115)
- 6. Global Change Master Directory Interchange Format
- 7. Climate and Forecast Convention...
- 8. Attribute Convention for Dataset Discovery
- 9. ISO8601...
- 10. GCMD Science Keywords
- 11. Climate and Forecast Standard Names
- 12. SeaDataNet
- 13. OPeNDAP

1.2. Scope

The intended audience of this document is technical staff at METNO responsible for implementing and maintaining software for metadata management in support of internal and external services including, but not restricted to:

- BarentsWatch
- Halo
- METSIS
 - Arctic Data Centre
 - WMO Global Cryosphere Watch
 - 0 ...

The purpose is to document datasets, not web services. Web services are provided through the data_access element.

1.3. Objectives

- 1. To facilitate documentation of data and products managed by METNO.
- 2. To facilitate metadata re-use between different projects and services at METNO.
- 3. To be compatible with the GCMD DIF and ISO19115/ISO19139 metadata standards as imposed by WMO and Norge Digitalt/INSPIRE.
- 4. To provide as lossless conversion between the different formats as possible.

Chapter 2. Metadata elements

2.1. metadata_identifier

Element	/mmd/metadata_identifier
Required	Yes
Repetition allowed	No
Description	Unique identifier for the dataset described by the metadata document. This identifier is used to identify a dataset across different systems.
	The MMD format does not make hard requirements on the format of the ids, but to be compatible with DIF the following characters are not allowed:
	• Backward Slash '\'
	• Forward Slash '/'
	• Colon ':'
	• Whitespaces ' '
	The identifier used currently is UUID.
Example XML	<metadata_identifier> 9663fc67-5687-4bf2-a274-f3826e41fdc8 </metadata_identifier>
DIF equivalent	/DIF/Entry_ID
ISO equivalent	/gmd:MD_Metadata/gmd:fileIdentifier/gco:CharacterString
MM2 equivalent	NA

2.2. last_metadata_update

Element	/mmd/last_metadata_update
Attributes	None
Required	Yes
Repetition Allowed	
Description	The last update of the metadata record.

Element	/mmd/last_metadata_update
Example XML:	<pre><last_metadata_update> 2012-10-31 </last_metadata_update></pre>
DIF equivalent	/DIF/Last_DIF_Revision_Date
ISO equivalent	/gmd:MD_Metadata/gmd:dateStamp This mapping is somewhat uncertain since the element is described as: "date that metadata was created." But GeoNetwork automatically updates on Save, so we take this to mean that it is last updated time for the metadata.
MM2 equivalent	Fetched from the associated .xmd file: /dataset@datestamp

2.3. metadata_status

Element	/mmd/metadata_status
Attributes	None
Required	Yes
Repetition allowed	No
Description	Status for the metadata record. This is configuration metadata and should not be misinterpreted as dataset_production_status . The only purpose of this tag is to determine whether the dataset should be indexed or not.
Example XML	<metadata_status> Active </metadata_status>
DIF equivalent	NA
ISO equivalent	NA
MM2 equivalent	This was not in MM2, but in the XMD-files.

2.4. collection

Element	/mmd/collection
Attributes	None
Required	Yes
Repetition allowed	Yes

Element	/mmd/collection
Description	The purpose of this tag is the same as for the ownertag in XMD files. It is used to identify which collection a dataset belong to. This is used to identify sets when serving metadata through e.g. OAI-PMH or to identify which data to present in e.g. a project specific portal when all metadata records are in the same repository. The keyword used to identify the collection should be short (e.g. NMDC, NMAP, SIOS,). See Collection keywords for details.
Example XML	<collection> NMDC </collection>
DIF equivalent	NA
ISO equivalent	NA
MM2 equivalent	ownertag (XMD)

2.5. title

Element	/mmd/title
Attributes	• xml:lang ISO language code for the language that the title is in
Required	Yes
Repetition allowed	Yes, but each repetition should have a different language.
Description	The title of the dataset. The language in the title is specificed in the xml:lang attribute. To be compatible with DIF the title cannot be longer than 220 characters.
Example XML	<title xml:lang="DenD"> OSISAF Northern Hemisphere Ice edge </title>
DIF equivalent	/Entry_Title
ISO equivalent	/gmd:MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citation/gmd:title/gco:CharacterString
MM2 equivalent	/metadata@name='title'

2.6. abstract

Element	/mmd/abstract
Attributes	• xml:lang ISO language code for the language that the abstract is in
Required	Yes
Repetition allowed	Yes, but each repetition should have a different language.

Element /mmd/abstract

Description The abstract should summarize and described the dataset.

The following guidelines for a good abstract follows (from DIF):

- Capitalization should follow standard constructs. For readability, all capital letters or all lower case letters should not be used. Use the appropriate case where applicable.
- Acronyms should be expanded to provide understanding.
- Where applicable, the abstract should also include brief statements on the following information:
- Data processing information (gridded, binned, swath, raw, algorithms used, necessary ancillary data sets).
- Date available.
- Data set organization (description of how data are organized within and by file).
- Scientific methodology or analytical tools.
- Time gaps in data set coverage.
- Units and unit resolution.
- Similarities and differences of these data to other closely-related data sets.
- Other pertinent information.
- Capitalization should follow standard constructs. For readability, all capital letters or all lower case letters should not be used. Use the appropriate case where applicable.
- Acronyms should be expanded to provide understanding.
- Where applicable, the abstract should also include brief statements on the following information:
- Data processing information (gridded, binned, swath, raw, algorithms used, necessary ancillary data sets).
- Date available.
- Data set organization (description of how data are organized within and by file).
- Scientific methodology or analytical tools.
- Time gaps in data set coverage.
- Units and unit resolution.
- Similarities and differences of these data to other closely-related data sets.
- Other pertinent information.

Element	/mmd/abstract
Example XML:	<pre><abstract xml:lang="en"> The daily analysis of sea ice concentration are obtained from operation satellite images of the polar regions. It is based on atmospherically corretected signal and an optimal sea ice concentration algorithm. This product is available for free from the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF). </abstract></pre>
DIF equivalent	/Summary/Abstract
ISO equivalent	/gmd:MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:abstract/gco:Characte rString
MM2 equivalent	/metadata@name='abstract'

2.7. temporal_extent

Element	/mmd/temporal_extent
Attributes	None
Required	Yes
Repetition allowed	Yes. Repetition is used when there are gaps in the dataset.
Description	Defines the temporal extent or coverage of the dataset. The extent is defined used two sub elements: • start_date: The start date for data collection or model coverage.
	 end_date: The end date for data collection or model coverage.
	If the dataset is not complete, the end_date element can be left empty.
Example XML:	<temporal_extent> <start_date>2012-01-01<start_date> <end_date>2012-02-01</end_date> </start_date></start_date></temporal_extent>
DIF equivalent	/DIF/Temporal_Coverage
ISO equivalent	/gmd:MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:extent/gmd:EX_Extent /gmd:temporalElement/gmd:EX_TemporalExtent/gmd:extent

Element	/mmd/temporal_extent
MM2 equivalent	/metadata@name='datacollection_period_from'
	and /metadata@name='datacollection_period_to

2.8. geographic_extent/rectangle

Element	/mmd/geographic_extent/rectangle
Attributes	srsName
Required	Yes
Repetition allowed	No
Description	The geographic extend of the datasets defined a rectangle in lat/lon projection. The extent is defined using the following child elements: • north: The northernmost point covered by the dataset. • south: The southernmost point covered by the dataset. • west: The westernmost point covered by the dataset. • east: The easternmost point covered by the dataset.
Example XML:	
DIF equivalent	/DIF/Spatial_Coverage
ISO equivalent	/gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:e xtent/gmd:EX_Extent/gmd:geographicElement/gmd:EX_GeographicBoundingB ox
MM2 equivalent	/metadata@name='bounding_box' For MM2 this is comma separated list so it needs to be split during conversion.

2.9. geographic_extent/polygon

Element	/mmd/geographic_extent/polygon
Attributes	None

Element	/mmd/geographic_extent/polygon
Required	No
Repetition allowed	No
Description	Geographic extent of the dataset described as a GML polygon. The projection of the coordinates can be specified in compliance with GML, but for maximum compatibility with internal systems all coordinates should be given in EPSG:4326.
Example XML:	<pre><polygon> <gml:polygon id="polygon" srsname="EPSG:4326"> <gml:exterior> <gml:linearring> <gml:pos> -180 -90 </gml:pos> <gml:pos> <gml:pos> 180 90 </gml:pos> -180 90 </gml:pos> -180 90 -180 90 -180 90 </gml:linearring></gml:exterior></gml:polygon></polygon></pre>
DIF equivalent	NA
ISO equivalent	/gmd:MD_Metadata/ gmd:identificationInfo/gmd:MD_DataIdentification/gmd:extent/gmd:EX_Extent /gmd:geographicElement/gmd:EX_BoundingPolygon/gmd:polygon/
MM2 equivalent	A similar field is found in the associated .xmd file under datasetRegion/lonLatPoints but these points are too many for efficient metadata handling and cannot be used directly.

2.10. location

Element	/mmd/location
Attributes	None
Required	No
Repetition allowed	No
Description	A textual identification of the area covered by the dataset.
	Describes the name of a place on Earth, a location within the Earth, a vertical location, or a location outside of Earth. The <location>* *keywords are a 5-level hierarchy of controlled keywords. This element follows the same format as DIF, but also allows other vocabularies as long as they have the same format the DIF vocabulary.</location>
	location_vocabulary: The name of the vocabulary that is used for locations. This element exists to allow different vocabularies to be used if the vocabulary provided by GCMD is not sufficient
	location_reference: URL to the authoritative source, preferably in SKOS/RDF.
	• location_category: The main category of the location name.
	 location_type: The main type within a specific location category
	 location_subregion1: Subregion1 keyword
	• location_subregion2: Subregion2 keyword
	• location_subregion3: Subregion3 keyword
	• detailed_location: Free text describing the location further
	Vocabularies used should have a well defined description of the area names, including geographical boundaries (e.g. like SeaDataNet and ICES have).
Example XML:	<pre><location> <location_vocabulary>gcmd</location_vocabulary> <location_category>CONTINENT</location_category> <location_type>NORTH AMERICA</location_type> <location_subregion1>UNITED STATES OF AMERICA</location_subregion1> <location_subregion2>MARYLAND</location_subregion2> <location_subregion3>BALTIMORE</location_subregion3> <detailed_location>Middle River</detailed_location> </location></pre>
DIF equivalent	/Location
ISO equivalent	NA
MM2 equivalent	There is no good mapping to MM2 even though the element /metadata@name='area' is somewhat similar.

2.11. dataset_production_status

Element	/mmd/dataset_production_status
Attributes	None
Required	Yes
Repetition allowed	No
Description	Production status for the dataset using a controlled vocabulary. The valid keywords are listed in Dataset production status. If set as "In Work", remember that end_date in temporal_extent can (should) be empty.
Example XML	<pre><dataset_production_status> In Work </dataset_production_status></pre>
DIF equivalent	/DIF/Data_Set_Progress The defined statuses are the same as in DIF so no mapping is required.
ISO equivalent	/gmd:MD_Metadata/gmd:status
MM2 equivalent	NA

2.12. dataset_language

Element	/mmd/dataset_language
Attributes	None
Required	No
Repetition allowed	No
Description	The language used in production, storage etc. of the dataset. The default for all datasets is English.
Example XML:	<dataset_language> en </dataset_language>
DIF equivalent	/DIF/Data_Set_Language
ISO equivalent	/gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:language/gmd:LanguageCode
MM2 equivalent	NA

2.13. operational_status

Element	/mmd/operational_status
Attributes	None
Required	No
Repetition allowed	No
Description	The current operational status of the product. Valid keywords are listed in Operational status.
Example XML:	<pre><operational_status> pre-operational </operational_status></pre>
DIF equivalent	NA
ISO equivalent	NA
MM2 equivalent	operational_status

2.14. access_constraint

Element	/mmd/access_constraint
Attributes	None
Required	No
Repetition allowed	No
Description	Limitations on the access to the dataset. See Access constraints for a list of valid values.
Example XML:	<access_constraint>Open for all</access_constraint>
DIF equivalent	/DIF/Access_Constraints
ISO equivalent	/gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:resourceConstraints/ gmd:MD_LegalConstraints/gmd:accessConstraints/gmd:MD_RestrictionCode = 'otherConstraints'
	/gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:resourceConstraints/ gmd:MD_LegalConstraints/gmd:otherConstraints/gco:CharacterString

Element	/mmd/access_constraint
MM2 equivalent	/metadata@name='distribution_statement'

2.15. use_constraint

Element	/mmd/use_constraint
Attributes	None
Required	No
Repetition allowed	No
Description	Restrictions on the use of the dataset.
	See Use constraintsfor a list of valid values.
Example XML:	<use_constraint> CC-BY </use_constraint>
DIF equivalent	/DIF/Use_Constraint
ISO equivalent	/gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:resourceConstraints/ gmd:MD_LegalConstraints/gmd:useLimitation/gco:CharacterString
MM2 equivalent	NA

2.16. personnel

Element	/mmd/personnel
Attributes	None
Required	Yes (see details on roles below)
Repetition allowed	Yes

Element /mmd/personnel

Description Relevant contact persons for the dataset. All datasets must have at least one entry of personnel specifying the Principal Investigator (role Investigator). The contact has the following sub elements

- role: The role the person has related to this dataset. The value must * come from Contact roles.
- name: The full name of the contact.
- email: The email address to the contact.
- phone: The phone number to the contact. The number should include land codes.
- fax: The fax number where the contact can receive faxes.
- organisation: The name of the organisation where the person is employed.
- contact_address contain the address information of the person. It consists of:
- address is the organization name, department, mail stop, street address, etc. of the person.
- city: is the city or town of the person.
- province_or_state: is the province, region or state.
- postal_code: is the postal code of the person.
- country: is the country of the person.

Only role, name and email are mandatory.

Example XML:

```
<personnel>
     <role>Technical contact</role>
     <name>Ole Dole</name>
     <email>ole.dole@example.com</email>
     <phone>004711111111</phone>
     <fax/>
     <contact_address>
     <address>Meteorologisk institutt, Henrik Mohnsplass
1</address>
     <city>0slo</city>
     cor_state>0slo/province_or_state>
     <postal code>0000</postal code>
     <country>Norway</country>
     </contact_address>
<personnel>
```

Element	/mmd/personnel
DIF equivalent	/Personnel There is a slight mismatch with DIF since DIF has first name, middle name and last name as separate elements. In export to DIF the name element is exported to <last_name> since this element is required.</last_name>
ISO equivalent	For the personell with the role "Metadata author" we place it in /gmd:MD_Metadata/gmd:contact/gmd:CI_ResponsibleParty since the ISO standard states that the this element should be "party responsible for the metadata information" For the other contacts we place them in /gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:pointOfContact
MM2 equivalent	/metadata@name='PI_name', only partly mapped.

2.17. data_center

Element	/mmd/data_center
Attributes	None
Required	No
Repetition allowed	Yes
Description	Description about the datacenter responsible for the distribution of the dataset. The element has the following child elements: • data_center_name: The name of the data center split into two child elements; short_name and long_name. • data_center_url: URL to the data center's main website. • dataset_id: The id assigned to the dataset by the data center. This may be different from metadata_identifier. Information on the data centre contact person for a data centre is conveyed through the element personnel. See personnel for more information. The role of a data center contact must be "Data center contact".

```
Element /mmd/data_center
  Example XML
                  <data_center>
                      <data_center_name>
                          <short_name>METNO</short_name>
                          <long_name>
                          Norwegian Meteorological Institute
                          </long_name>
                      </data_center_name>
                      <data center url>http://met.no</data center url>
                      <dataset_id>45dlkf-234-df</dataset_id>
                      <contact>
                          <role>Data center contact</role>
                          <name>Ole Dole</name>
                          <email>ole.dole@example.com</email>
                          <phone>004711111111</phone>
                          <fax/>
                      <contact>
                 </data_center>
 DIF equivalent /DIF/Data_Center
 ISO equivalent /gmd:MD_Metadata
               /gmd:distributionInfo/gmd:MD_Distribution/gmd:distributor/gmd:MD_Distribu
               tor/gmd:distributorContact/gmd:CI_ResponsibleParty
MM2 equivalent NA
```

2.18. data_access

Element	/mmd/data_access
Attributes	None
Required	No
Repetition allowed	Yes

Element	/mmd/data_access
Description	Description of how to access the data in the dataset. This element has the following child elements:
	• type: Description of the type of data access. Data Access Types
	• name: The name of the data access mechanism.
	 description: A short description of the data access type and purpose (if necessary).
	• resource: The URL used to access the data. E.g. for WMS this should point at the GetCapabilities document.
	• wms_layers: WMS servers behave very different. Some servers has huge GetCapabilities documents with several thousand entries. The purpose of this element is to predefine a number of WMS layers. The first WMS layer will be presented initially1 if the user requests visualisation.
Example XML:	<pre><data_access></data_access></pre>
DIF equivalent	/Related_UrlRL
ISO equivalent	/gmd:MD_Metadata/ gmd:distributionInfo/gmd:MD_Distribution/gmd:transferOptions/gmd:MD_Digi talTransferOptions/gmd:onLine/gmd:CI_OnlineResource
MM2 equivalent	/metadata@name='dataref'
	/metadata@name='dataref_OPENDAP'
	/metadata@name='dataref_WMS'
	WMS layer information is currently available only through XMD documents.

2.19. related_dataset

Element	/mmd/related_dataset
Attributes	 relation_type. Valid content is shown below: parent: This dataset is a child dataset of the referenced dataset. i.e. the other dataset is its parent. auxiliary: This dataset is auxiliary data for the referenced dataset.
Required	No
Repetition allowed	Yes.
Description	Specifies the relation between this dataset and another dataset. The type of relationship is determined by the relation_type attribute. The content of the element is a valid identifier as identified in the metadata_identifier element.
Example XML	<related_dataset relation_type="OparentO"> 9663fc67-5687-4bf2-a274-f3826e41fdc8 </related_dataset>
DIF equivalent	NA
ISO equivalent	If the relation_type = "parent". In other cases there are is no mapping. /gmd:MD_Metadata/gmd:parentIdentifier
MM2 equivalent	Within MM2 this was determined by a directory structure and not the content of the metadata files.NA

2.20. related_information

Element	/mmd/related_information
Attributes	None
Required	No
Repetition allowed	Yes

Element	/mmd/related_information
Description	 Description of how to access the data in the dataset. This element has the following child elements: type: Description of the type of information. A controlled vocabulary is required. See Related Information types. resource: The URL used. description: Textual description on the resource. This element should be used for e.g. THREDDS landing pages. i.e. that is the pages providing the overview of services (HTTP, OPeNDAP, WMS,) for a dataset. The specific web services for data access are to be added using data_access.
Example XML:	<pre><related_information></related_information></pre>
DIF equivalent	/Related_Url View Related Information
ISO equivalent	/gmd:MD_Metadata/ gmd:distributionInfo/gmd:MD_Distribution/gmd:transferOptions/gmd:MD_Digi talTransferOptions/gmd:onLine/gmd:CI_OnlineResource
MM2 equivalent	/metadata@name='dataref'

2.21. iso_topic_category

Element	/mmd/iso_topic_category
Attributes	none
Required	Yes
Repetition allowed	Yes
Description	ISO topic category fetched from a controlled vocabulary. Valid keywords are listed in ISO Topic categories. This field is required for compatibility with DIF and ISO

Element	/mmd/iso_topic_category
Example XML	<iso_topic_category>ocean</iso_topic_category>
DIF equivalent	/DIF/ISO_TOPIC_Category
ISO equivalent	/gmd:MD_Metadata/gmd:identificationInfo/ gmd:MD_DataIdentification/gmd:topicCategory/gmd:MD_TopicCategoryCode
MM2 equivalent	/MM2/metadata[@name='topiccategory']

2.22. keywords

Element	/mmd/keywords
Attributes	• vocabulary: The vocabulary where the keyword is fetched from. The vocabulary "none" is used if the keyword is not from a vocabulary. The recommended vocabulary to use is "GCMD", see Variable/parameter descriptions for further details. Other valid vocabularies are e.g. "WIGOS". Search services are currently based on GCMD Science keywords.
Required	Yes
Repetition allowed	Yes. It is expected that different keyword elements have different vocabulary child elements.
Description	A set of keywords describing the dataset. The keywords should be fetched from a pre-defined vocabulary, but this is not a requirement. This element has two child elements:
	• reference: URI to machine readable form of the vocabulary used.
	• keyword: A single keyword describing the dataset. This can be hierarchically like GCMD using a separator.
	• separator: the character used as separator. This could be e.g. the GCMD
	• '>', but in XML this is encoded as '>'. Other valid separators are: / Valid controlled vocabularies are GCMD Science Keywords [10], Climate and Forecast Standard Names [11], SeaDataNet [12] etc. The primary usage for this element is to describe variables in the dataset, but it could also be used to describe other features.

Element	/mmd/keywords
Example XML:	<pre><keywords vocabulary="DMyOwnVocabD"></keywords></pre>
DIF equivalent	/DIF/Parameters
ISO equivalent	/gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords/ gmd:MD_Keywords/gmd:keyword/gco:CharacterString
MM2 equivalent	/metadata@name='variable'

2.23. project

Element	/mmd/project
Attributes	None
Required	No
Repetition allowed	Yes
Description	Project where the dataset was generated or collected. The element is described using two child element: • short_name: is the abbreviated name of the project from which the data were collected. • long_name: is the full name of the project from which the data were collected. These names must be maintained in a controlled vocabulary.
Example XML:	<pre><pre><pre><pre><pre><pre>< cshort_name>ICE</pre></pre></pre></pre></pre></pre>

Element	/mmd/project
DIF equivalent	/DIF/Project
ISO equivalent	NA
MM2 equivalent	/metadata@name='project_name' (usually mapped to /project/long_name)

2.24. platform

Element	/mmd/platform
Attributes	None
Required	No
Repetition allowed	Yes
Description	The platform used to collect the data. The element is described using two child elements:
	 short_name is the abbreviated name of the platform used to acquire the data.
	• long_name is the full name of the platform used to acquire the data.
	• This element is used to store information for NBS products (Sentinel platforms).
	See platform for list.
Example XML:	<pre><platform></platform></pre>
DIF equivalent	/DIF/Platform
ISO equivalent	For ISO we translate platforms to keywords from a specific vocabulary. /gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords /gmd:MD_Keywords/gmd:keyword/gco:CharacterString
MM2 equivalent	

2.25. instrument

Element	/mmd/instrument
Attributes	None
Required	No
Repetition allowed	Yes
Description	Instrument used to collect the data. The element is described using two child elements:
	 short_name: is the abbreviated name of the instrument used to acquire the data.
	• long_name: is the full name of the instrument used to acquire the data.
	 mode: is the operating mode of the instrument (if applicable).
	• polarisation: polarisation mode when applicable.
	 product_type: is the type of product generated (if applicable)
	See Instruments, Instrument modes and Polarisation modes for list.
Example XML:	<pre><instrument></instrument></pre>
DIF equivalent	/DIF/Instrument
	For ISO we just translate instruments to keywords belonging to a specific vocabulary.
	/gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords /gmd:MD_Keywords/gmd:keyword/gco:CharacterString

2.26. activity_type

Element	/mmd/activity_type
Attributes	None
Required	No
Repetition allowed	Yes.
Description	The activity used to collect the data. Valid keywords are listed in Activity type.
Example XML:	<activity_type>Space Borne Instrument</activity_type>
DIF equivalent	NA
ISO equivalent	NA
MM2 equivalent	/metadata@name='activity_type'

2.27. reference

Element	/mmd/reference
Attributes	Type. Attribute used to signal if the reference is specified as freetext or using fields. Authors note: This attribute is required for easier identification of different types references based using Xpath.:
Required	No
Repetition allowed	Yes

Element /mmd/reference

Description The reference element describes key bibliographic citations pertaining to the data set.

> Bibliographic citations may be populated as a block of text in the free-text field or as text in subfields (see below). Citations in the free-text field may be provided in styles used by professional scientific journals such as the American Geophysical Union (AGU), American Psychological Association (APA), or Modern Language Association (MLA).

The field can contain the following sub elements:

- author: The name of the individual(s) or organization(s) with primary intellectual responsibility for the cited resource's development. This is required if this form is used.
- publication_date: The publication/reference date for the cited resource. If the Publication Date is not known, please use "Unknown". This is required if this form is used.
- title: The title (name) of the cited resource. This is required if this form is used.
- Series: The name of a resource series, or aggregate (collection) resource of which the cited resource is a part.
- edition: The version of a cited resource.
- volume: The order or sequence of a resource in a series or set.
- issue: The issue number of a resource (usually within a volume).
- report number: The unique number or code assigned to a resource by the issuing organization(s).
- publication_place: The name of the city (and state or province and country if needed) where the resource was made available. This is required if this form is used.
- publisher: The name of the individual(s) or organization(s) that made the resource available. This is required if this form is used.
- pages: The range of pages or total number of pages of a cited resource.
- isbn: International Standard Book Number.
- doi: Digital Object Identifier.
- online_resource: The URL of the online resource containing the cited resource.
- other_reference_details: Additional free-text reference information.

Element	/mmd/reference
Example XML	<pre><reference type="OfieldsO"> +</reference></pre>
DIF equivalent	/Reference
ISO equivalent	Not known if this maps to any element in ISO.
MM2 equivalent	NA

2.28. dataset_citation

Element	/mmd/dataset_citation
Attributes	None
Required	No
Repetition allowed	Yes

Element /mmd/dataset_citation

Description A description of how the dataset should be properly cited. The element has the following child elements:

- dataset_creator: The name of the organization(s) or individual(s) with primary intellectual responsibility for the data set's development.
- dataset_editor: Dataset Editor is the individual(s) responsible for the processing or re-processing of a particular dataset.
- dataset_title: The title of the data set; this may be the same as Entry Title.
- dataset_series_name: The name of the dataset series, or aggregate dataset of which the dataset is a part.
- dataset_release_date: The date when the data set was made available for
- dataset_release_place: The name of the city (and state or province and country if needed) where the data set was made available for release.
- dataset_publisher The name of the individual or organization that made the data set available for release.
- Version: The version of the data set.
- issue identification: The volume or issue number of the publication (if applicable).
- data_presentation_form: The mode in which the data are represented, e.g. atlas, image, profile, text, etc.
- other citation details: Additional free-text citation information. For <other citation details>NOAA/NCDC example, Paleoclimatology Program</other_citation_details>
- dataset_doi: The Digital Object Identifier (DOI) of a data set.

Element /mmd/dataset_citation Example XML: <dataset_citation> <dataset_creator>OMI SIPS</dataset_creator> <dataset editor>J. Smith, D. Johnson, M. Davis</dataset editor> <dataset_title>OMI/Aura Level 1B VIS Global Geolocated Earth Shine Radiances 1-orbit L2 Swath 13x24 km</dataset title> <dataset series name>OML1BRVG</dataset series name> <dataset_release_date>2009</dataset_release_date> <dataset release place>NASA Goddard Space Flight Center</dataset release place> <dataset_publisher>U.S. Geological Survey</dataset_publisher> <version>004</version> <data_presentation_form>Digital Science Data</data_presentation_form> <online_resource>http://disc.gsfc.nasa.gov/Aura/OMI/oml1brvg_v004.s html</online_resource> </dataset citation> DIF equivalent /DIF/Data_Set_Citation ISO equivalent There are no direct equivalent in ISO for this element as a whole, but parts can be put into citation information. dataset release date can be translated to /gmd:MD_Metadata /gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citati on/gmd:date/gmd:CI_Date/gmd:date/gco:DateTime where gmd:CI DateTypeCode has attribute codeListValue="publication" **version** can be translated to gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:cit ation/gmd:CI_Citation/gmd:edition other citation details can be translated to gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:cit ation/gmd:CI_Citation/gmd:otherCitationDetails MM2 equivalent NA

2.29. cloud_cover

Element	/mmd/cloud_cover
Attributes	

Element	/mmd/cloud_cover
Required	No
Repetition allowed	No
Description	This is an element specifically made for the NBS project and initially in support for handling Sentinel-2 data. It is however also relevant for other optical satellite imagery whether AVHRR (NPOES, EPS), VIIRS (NPP), MERIS (ENVISAT) or MODIS (EOS). It describes the overall cloud cover in the scene and it only relates to the number of valid pixels in the scene (excluding missing values for map projected scenes). The element has one sub element: value: The actual cloud cover in percentage of the valid pixels. This is indicated with one attribute.
Example XML	<pre><cloud_cover> <value unit="0%0">45</value> </cloud_cover></pre>
DIF equivalent	NA
ISO equivalent	NA
MM2 equivalent	NA

2.30. scene_cover

Element	/mmd/scene_cover
Attributes	
Required	No
Repetition allowed	No

Element	/mmd/scene_cover
Description	This is an element specifically made for the NBS project, but based on previous experience with AVHRRm and initially in support for handling Sentinel-2 data. It is however also relevant for other optical satellite imagery whether AVHRR (NPOES, EPS), VIIRS (NPP), MERIS (ENVISAT) or MODIS (EOS). It describes the overall coverage of valid data within a scene. This is to let the user know whether there is only a very limited number of valid data points in a scene. The reason for this is the map projected scenes where you may only have a small fraction of the predefined tile covered by the satellite. It relates to the total number of pixels in the scene. The element has one sub element: • value: The actual cloud cover in percentage of the valid pixels. This is indicated with one attribute.
Example XML	<scene_cover> <value unit="0%0">45</value> </scene_cover>
DIF equivalent	NA
ISO equivalent	NA
MM2 equivalent	NA

Chapter 3. Explanations

3.1. Interpretation of child and parent relationship

If a dataset has a <related_dataset relation_type="parent"> element it means that the referred dataset is this datasets parent. This relationship does not impose any specific characteristics on either the parent or the child dataset. For instance it is not required that the parents dataset is a dataset series and that the child dataset is a dataset within the series. The relationship only causes the child dataset to inherit elements from the parent dataset. Exact semantics for inheritance is to be decided.

Chapter 4. Controlled vocabularies

4.1. Collection keywords

CC	Datasets related to the Cryosphere in the context of the CryoClim project.
	Remote sensing datasets or higher order products based on remote sensing
144411	datasets generated in the context of the Norwegian Satellite Earth Observation Database for Marine and Polar Research (NORMAP).
ADC	Datasets related to the Arctic Data Centre. This is the fall back collection, but it might be that some datasets are not officially part of the ADC collection and thus should only be searchable, but not exported using OAI-PMH.
GCW	Datasets related to the WMO Global Cryosphere Watch.
NMDC	Datasets relevant for the Norwegian Marine Data Centre. This is basically datasets of a marine origin, regardless of whether they are of operational or scientific origin.
SIOS	Datasets related to the Svalbard Integrated Arctic Earth Observing System.
NSDN	Datasets related to the Norwegian Scientific Data Network (NorDataNet).
DOKI	Datasets related to the national contribution to the International Polar Year.
DAM	Datasets in the context of EU FP6 project DAMOCLES:
ACCESS	Datasets in the context of EU FP7 project ACCESS.
NBS	Datasets in the context of National Ground Segment for satellite data. This is primarily Sentinel datasets.
APPL	Datasets in the context of EU H2020 project APPLICATE.
YOPP	Datasets in the context of YOPP.

collection

4.2. Dataset production status

Code	Description
Planned	Refers to data sets to be collected in the future and are thus unavailable at the present time. For Example: The Hydro spacecraft has not been launched, but information on planned data sets may be available.
In Work	Refers to data sets currently undergoing production or data that is continuously being collected or updated. For Example: data from the AIRS instrument on Aqua is being collected continuously.
Complete	Refers to data sets in which no updates or further data collection will be made. For Example: Nimbus-7 SMMR data collection has been completed.
Obsolete	A new version of the dataset has been generated. The new version should be used, this is kept for back tracing.

4.3. ISO to MMD mapping

ISO Code	MMD code
completed	Complete
historicalArchive	Complete
obsolete	Obsolete
onGoing	In Work
planned	Planned
required	Planned
underDevelopment	Planned

4.4. MMD to ISO mapping

MMD Code	ISO code
Planned	Planned
In Work	onGoing
Complete	Complete
Obsolete	obsolete

4.5. Operational status

Code	Description
Operational	This is used to tag datasets that are continously being evaluated and validated. The production of the data is under the responsibility of an authority and follows a review and documentation process. Operational data are generated on a continuous basis with a temporal perspective.
Pre-Operational	This is used to describe data that are under review for becoming the next operational delivery. It is usually produced continuously, but not necessarily distributed to users not specifically. It covers the intermediate step between Experimental and Operational.
Experimental	This is used to describe data that belongs to a development and production chain that provides operational data. It is the first step in the process of becoming operational.
Scientific	This is used to describe purely scientific products. That is products generated through scientific projects and usually with a limited temporal perspective.

operational_status

4.6. Access constraints

Code	Description
Open	These data are freely available to everyone and can be made available without any restrictions.
only (automated	These data are available for users as long as they register with name, affiliation, and a valid email address. The verification process can be automated.
	These data are available for users within a restricted community. This community determines the authorization mechansim to utilise whether this being IP-address, community specific users names or other.
	Information on these data MUST NOT be exposed externally. This implies that neither metadata can be exposed externally.

access_constraint

4.7. Use constraints

Code	Description
Public domain	All rights for these data are waived. This should be relevant when data are released using CCO, Open Data Commons PDDL or similar.
Attribution	These data may be freely copied, shared, used, modified and transformed as long as you attribute the owner/creator of the original dataset. This is used for Creative Commons Attribution.
Share-alike	This is a bit stronger than attribution as it requires any work based upon these data to be shared using the same conditions as the original data. It implies attribution and limits the freedom for new data generated.
Noncommercial	Data may be used for noncommercial purposes only. This implies share-alike (and attribution).
No derivative	FIXME

use_constraint

4.8. Activity type

Controlled vocabulary used to describe activity types. Rather than using the term observation type or platform which possibly could describe the nature of observed datasets, activity type is used to filter between both observations and simulations that possibly are describing the same phenomena. Activity types are used to identify the origin of the dataset documented within

METAMOD. This is not an identification of the observation platform (e.g. specific vessel, SYNOP station or satellite), but more the nature of the generation process (e.g. simulation, in situ observation, remote sensing etc). It is useful in the context of filtering data when searching for relevant datasets.

Code	Description
Aircraft	Observations made during a flight trajectory. The observations caninclude remote sensing instruments, dropsondes or in situ measurements. Both manned and unmanned vehicles are covered by this term. The outputis typically a trajectory, but could also be profiles or points.
Space Borne Instrument	Observations or analysed products based upon data from a space borneinstrument (typically onboard a satellite). The nature of the output is typically gridded of type imagery or profiles.
Numerical Simulation	Data are generated by the use of a numerical simulation of theatmosphere, the ocean, the climate or similar. Statistical analysis is not covered by this.
Climate Indicator	This indicates a dataset that has been generated by analysis of somedata with the emphasis on being representative in a climate context (e.g. consistent in time). Furthermore, a climate indicator is a "compact" representation of the feature studied (e.g. the temporal evolution of area covered by sea ice in the Arctic). Climate Indicatorsare frequently linked to GCOS requirements.
station(Land station) (Field	This is used to tag datasets generated from a site located on land. Thiscan be a permanent (e.g. a SYNOP or TEMP station) or a temporary site (e.g. a field experiment).
_	This is used to identify datasets generated during cruises. Typically it describes a full dataset generated in a context, possibly describing both ocean and atmospheric conditions.
	This is used to describe ocean stations that are fixed in space. Typically this is moorings, anchored buoys, oil rigs etc.
	This is used to describe ocean stations that are moving around. Typically this is gliders and drifting buoys.
	This is used to tag datasets generated from a site located on driftingsea ice or some other ice sheet (possibly on land but moving). It typically describes a temporary site (e.g. a field experiment). It would also be used to describe ships frozen in ice and drifting e.g. across the Arctic as well as Ice Thethered Platforms (ITP) and Ice Mass BalanceBuoys (IMBB).
_	This is not much used within environmental science, but comes in usefulsometimes. It is used to cover the results of interviews and questionnaires especially in interdisciplinary science.

Code	Description
	This is used to tag datasets containing imagery or PDF documents. This could e.g. be a time lapse photographic session of a specific site illustrating e.g. snow cover or cloud cover. It can also be used to tag documents or maps describing the nature of a field station. It would then require datasets to be linked (which currently is not supported).

activity_type

4.9. Variable/parameter descriptions

For description of parameters MMD is currently relying on GCMD Science Keywords. The GCMD Science Keywords are available in multiple forms.

GCMD Science Keywords

- https://gcmdservices.gsfc.nasa.gov/kms/concepts/concept_scheme/sciencekeywords/?format=csv comma separated file
- 2. https://gcmdservices.gsfc.nasa.gov/kms/concepts/concept_scheme/sciencekeywords/? format=json JSON
- 3. https://gcmdservices.gsfc.nasa.gov/kms/concepts/concept_scheme/sciencekeywords/?format=rdf RDF
- 4. https://gcmdservices.gsfc.nasa.gov/kms/capabilities?format=html API

However many datasets are encoded using the Climate and Forecast convention. In this CF Standard Names are used. These are available through http://cfconventions.org/Data/cf-standard-names/70/build/cf-standard-name-table.html and can be converted to GCMD Science Keywords using http://dap.onc.uvic.ca/erddap/convert/keywords.html (further information to be provided).

For observational data WMO Integrated Global Observing System (WIGOS) observed parameter descriptions can be used. These are available at https://codes.wmo.int/wmdr/. Use information from the sections on observed variable.

keywords

4.10. Platform

Code	Description
S1A	Sentinel-1A
S1B	Sentinel-1B
S2A	SENTINEL-2A
S2B	SENTINEL-2B
S3A	SENTINEL-3A

Code	Description
S3B	SENTINEL-3

platform

4.11. Instruments

CSAR	Synthetic Aperature Radar C-band	Relates to Sentinel-1.
MSI	Multi-Spectral Instrument	Relates to Sentinel-2. A push-broom optical sensor.
OLCI	Ocean and Land Colour Instrument	Relates to Sentinel-3. A push-broom imaging spectrometer instrument.
SLSTR	Sea and Land Surface Temperature Radiometer	Relates to Sentinel-3. A dual view (near- nadir and backward views) conical imaging radiometer.
MWR	Microwave Radiometer	Relates to Sentinel-3.

instrument

4.12. Instrument modes

Code	Description	SM
_	Relates to Sentinel-1	IW
		EW
	Relates to Sentinel-1	WV

instrument

4.13. Polarisation modes

Code	Description
НН	Horisontally transmitted and Horisontally received
VV	Vertically transmitted and Horisontally received
HH+HV	
VV+VH	

Code	Description
HV+HH	
VH+VV	

instrument

4.14. Product type

Code	Content	Description
SLC	Single Look Complex	Single Look Complex (SLC) products consist of focused SAR data, georeferenced using orbit and attitude data from the satellite, and provided in slant-range geometry.
GRD	Ground Range Detected	Ground Range Detected (GRD) products consist of focused SAR data that has been detected, multi-looked and projected to ground range using an Earth ellipsoid model such as WGS84.
OCN	Ocean	Ocean (OCN) products for wind, wave and currents applications derived from the SAR data.

instrument

4.15. Contact roles

Code	Description
Investigator	The person who headed the investigation or experiment that resulted in the acquisition of the data described (i.e., Principal Investigator, Experiment Team Leader) and knows the details on data collection and processing.
Technical contact	The person who is knowledgeable about the technical content of the data (quality, processing methods, units, available software for further processing)
Metadata author	The main responsible person for the generation of the metadata for this dataset. Other people could have been involved, but this is the main contact with regard to the metadata.

personnel

4.16. MMD to ISO mapping

MMD Code	ISO code
Investigator	principalInvestigator
Technical contact	pointOfContact
Metadata author	author

4.17. ISO Topic categories

Code	Description
farming	earing of animals or cultivation of plants. For example, resources describing irrigation, aquaculture, herding, and pests and diseases affecting crops and livestock.
biota	naturally occurring flora and fauna. For example, resources describing wildlife, biological sciences, ecology, wilderness, sea life, wetlands, and habitats.
boundaries	legal land descriptions.
	atmospheric processes and phenomena. For example, resources describing cloud cover, weather, atmospheric conditions, climate change, and precipitation.
economy	economic activities or employment. For example, resources describing labor, revenue, commerce, industry, tourism and ecotourism, forestry, fisheries, commercial or subsistence hunting, and exploration and exploitation of resources such as minerals, oil, and gas.
elevation	height above or below sea level. For example, resources describing altitude, bathymetry, digital elevation models, slope, and products derived from this information.
environment	environmental resources, protection, and conservation. For example, resources describing pollution, waste storage and treatment, environmental impact assessment, environmental risk, and nature reserves.
geoscientificinformation	earth sciences. For example, resources describing geophysical features and processes, minerals, the composition, structure and origin of the earth's rocks, earthquakes, volcanic activity, landslides, gravity information, soils, permafrost, hydrogeology, and erosion.
health	health services, human ecology, and safety. For example, resources describing human disease and illness, factors affecting health, hygiene, mental and physical health, substance abuse, and health services.
	base maps. For example, resources describing land cover, topographic maps, and classified and unclassified images.
intelligenceMilitary	military bases, structures, and activities. For example, resources describing barracks, training grounds, military transportation, and information collection.
inlandWaters	inland water features, drainage systems, and their characteristics. For example, resources describing rivers and glaciers, salt lakes, water use plans, dams, currents, floods, water quality, and hydrographic charts.

Code	Description
location	positional information and services. For example, resources describing addresses, geodetic networks, postal zones and services, control points, and place names.
oceans	features and characteristics of salt water bodies excluding inland waters. For example, resources describing tides, tidal waves, coastal information, and reefs.
planningCadastre	land use. For example, resources describing zoning maps, cadastral surveys, and land ownership.
society	characteristics of societies and cultures. For example, resources describing natural settlements, anthropology, archaeology, education, traditional beliefs, manners and customs, demographic data, crime and justice, recreational areas and activities, social impact assessments, and census information.
structure	man-made construction. For example, resources describing buildings, museums, churches, factories, housing, monuments, and towers.
transportation	means and aids for conveying people and goods. For example, resources describing roads, airports and airstrips, shipping routes, tunnels, nautical charts, vehicle or vessel location, aeronautical charts, and railways.
utilitiesCommunications	energy, water and waste systems, and communications infrastructure and services. For example, resources describing hydroelectricity, geothermal, solar, and nuclear sources of energy, water purification and distribution, sewage collection and disposal, electricity and gas distribution, data communication, telecommunication, radio, and communication networks.

iso_topic_category

4.18. Related Information types

Code	Description
Project home page	URI to the project home page generating the data.
Users guide	URI to a users guide or product manual for the dataset.
Dataset landing page	A dataset landing page.
Observation facility	For observational data (in situ or remote) an information page containing more detailed information on the observation facility following OGC Observations and Measurements or WMO Integrated Global Observing System approach.
Extended metadata	Additional unspecified metadata on the data.

related_information

4.19. Data Access Types

Code	Description
HTTP	Direct access to the full data file. May require authentication, but should point directly to the data file or a catalogue containing the data.
OPeNDAP	Open-source Project for a Network Data Access Protocol
OGC WMS	OGC Web Mapping Service, URI to GetCapabilities Document.
OGC WFS	OGC Web Feature Service, URI to GetCapabilities Document.
OGC WCS	OGC Web Coverage Service, URI to GetCapabilities Document.
FTP	File Transfer Protocol.
ODATA	Open Data Protocol.

data_access