Science Keywords

Element Description

The Science Keywords element allows relevant Earth science keywords to be associated with a dataset to better enable data search and discovery. Th e Science Keywords are chosen from a controlled keyword hierarchy maintained in the Keyword Management System (KMS). A list of valid Science Keywords can be found here: https://gcmd.earthdata.nasa.gov/kms/concepts/concept_scheme/sciencekeywords?format=csv.

(It is important to note that the "EARTH SCIENCE SERVICES" keywords included at the top of the list in the csv file should *not* be used as Science Keywords. Valid Science Keywords start with the "EARTH SCIENCE" Category).

Best Practices

Science keywords are important for the precise search and retrieval of data and should accurately represent the dataset being described. As a rule of thumb, science keywords should represent the scientific parameters being provided in the data as well as any broader conceptual terms that may aid in describing the data. At a minimum, one science keyword hierarchy must be provided, and this hierarchy must go down to the 'Term' level of detail. The 'Detailed Variable' element is the only science keyword element that is not controlled by the KMS. The 'Detailed Variable' keyword should only be used if there is a very specific parameter provided in the data which is not adequately described by keywords in the KMS. If a particular science keyword is missing from the KMS, it is possible to put in a request to have it added. The KMS is managed by the Global Change Master Directory (GCMD) and new keyword requests may be made through the GCMD Keywords Community Forum.

All positions in the science keyword hierarchy must be populated until the desired level of detail is reached. Skipping or leaving blank a position in the keyword hierarchy will render the keyword invalid. The only exception to this is the 'Detailed Variable' element; a Detailed Variable keyword may be provided as long as it is preceded by the required Category, Topic, and Term keywords. Science keywords are not case sensitive.

Examples:

ScienceKeywords/Category: "EARTH SCIENCE"

ScienceKeywords/Topic: "ATMOSPHERE"

ScienceKeywords/Term: "CLOUDS"

ScienceKeywords/VariableLevel1: "TROPOSPHERIC/LOW LEVEL CLOUDS (OBSERVED/ANALYZED)"

ScienceKeywords/VariableLevel2: "STRATOCUMULUS"

ScienceKeywords/VariableLevel3: "STRATOCUMULUS CUMILIFORMIS"

ScienceKeywords/DetailedVariable: "STRATOCUMULUS VESPERALIS"

ScienceKeywords/Category: "EARTH SCIENCE"

ScienceKeywords/Topic: "BIOSPHERE"

ScienceKeywords/Term: "ECOLOGICAL DYNAMICS"

ScienceKeywords/VariableLevel1: "FIRE ECOLOGY"

ScienceKeywords/VariableLevel2: "FIRE MODELS"

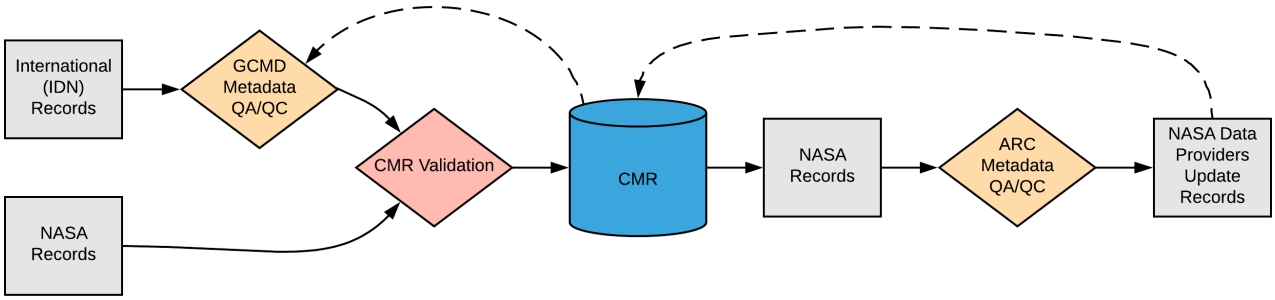
Element Specification

One Science Keyword is required and more Science Keywords may be provided (Cardinality: 1..*). For every science keyword that is provided the sub elements of Category, Topic, and Term are required.

Model	Element	Type	Usable Valid Values	Constraints	Required?	Cardinality
UMM-C	ScienceKeywords/Category	String	Science Category Keywords	KMS controlled	Yes	1
UMM-C	ScienceKeywords/Topic	String	Science Topic Keywords	KMS controlled	Yes	1
UMM-C	ScienceKeywords/Term	String	Science Term Keywords	KMS controlled	Yes	1
UMM-C	ScienceKeywords /VariableLevel1	String	Science Variable_Level_1 Keywords	KMS controlled	No	0..1
UMM-C	ScienceKeywords /VariableLevel2	String	Science Variable_Level_2 Keywords	KMS controlled	No	0..1
UMM-C	ScienceKeywords /VariableLevel3	String	Science Variable_Level_3 Keywords	KMS controlled	No	0..1
UMM-C	ScienceKeywords /DetailedVariable	String		1 - 80 characters (Uncontrolled/Free-Text)	No	0..1

Metadata Validation and QA/QC

All metadata entering the CMR goes through the below process to ensure metadata quality requirements are met. All records undergo CMR validation before entering the system. The process of QA/QC is slightly different for NASA and non-NASA data providers. Non-NASA providers include interagency and international data providers and are referred to as the International Directory Network (IDN).



Please see the expandable sections below for flowchart details.

- Manual Review
  - Identify errors, discrepancies, or omissions.
  - Verify that all pertinent keywords have been applied.
  - Verify that existing facets and other controlled keyword values are consistent and suitable for the data.
- Automated Review
  - Check that the field has been populated.
  - Check that the field is populated with a valid value from KMS.
  - Check that the field value is not a duplicate.
  - Check that the 'Detailed_Variable' field length is not greater than 80 characters.

- This element is required and at least 1 science keyword must exist.
- For every science keyword the sub elements of Category, Topic, and Term must exist.
- All science keyword sub-elements except for DetailedVariable must be valid according to the keyword management system. Currently the CMR issues a warning if this constraint is violated.

ARC Priority Matrix

Priority Categorization	Justification

Red = High Priority Finding	This element is categorized as highest priority when:
	• No Science Keywords are provided.
	• A Science Keyword does not align with the KMS.
	  ◦ The Science Keyword does not exist in the KMS.
	  ◦ A keyword(s) is missing from the keyword hierarchy.
	  ◦ A keyword(s) is placed in the incorrect position of the keyword hierarchy (e.g. a Variable Level 2 keyword is placed in the Variable Level 1 field).
	• A Science Keyword provided is not appropriate for the dataset.

Yellow = Medium Priority Finding	This element is categorized as medium priority when: • A recommendation is made to add a relevant Science Keyword to the metadata. • A recommendation is made to add to an existing keyword in the metadata (i.e. to extend a keyword hierarchy down to a more detailed level).
Blue = Low Priority Finding	Not Applicable
Green = No Findings /Issues	The element is provided and follows all applicable criteria specified in the best practices section above.

ARC Automated Checks

ARC uses the pyQuARC library for automated metadata checks. Please see the pyQuARC GitHub for more information.

Dialect Mappings

DIF 9 **(Note: DIF-9 is being phased out and will no longer be supported after 2018)**

DIF 10

Science_Keywords are required. An unlimited amount of Science Keywords may be provided (Cardinality: 1..*)

UMM-C Element	DIF 10 Path	Type	Constraints	Required in DIF 10?	Cardinality	Notes
ScienceKeywords /Category	Science_Keywords /Category	String	KMS controlled	Yes	1	The category keyword will always be "EARTH SCIENCE"
ScienceKeywords/Topic	Science_Keywords/Topic	String	KMS controlled	Yes	1	
ScienceKeywords/Term	Science_Keywords/Term	String	KMS controlled	Yes	1	
ScienceKeywords /VariableLevel1	Science_Keywords /Variable_Level_1	String	KMS controlled	No	0..1	
ScienceKeywords /VariableLevel2	Science_Keywords /Variable_Level_2	String	KMS controlled	No	0..1	
ScienceKeywords /VariableLevel3	Science_Keywords /Variable_Level_3	String	KMS controlled	No	0..1	
ScienceKeywords /DetailedVariable	Science_Keywords /Detailed_Variable	String	1 - 80 characters (Uncontrolled/Free-Text)	No	0..1	

Example Mapping

DIF 10

```
<Science_Keywords>
  <Category>EARTH SCIENCE</Category>
  <Topic>BIOSPHERE</Topic>
  <Term>VEGETATION</Term>
  <Variable_Level_1>VEGETATION INDEX<
/Variable_Level_1>
  <Variable_Level_2>NORMALIZED DIFFERENCE VEGETATION
INDEX (NDVI)</Variable_Level_2>
  <Detailed_Variable>0.9 DENSITY</Detailed_Variable>
</Science_Keywords>
<Science_Keywords>
  <Category>EARTH SCIENCE</Category>
  <Topic>BIOSPHERE</Topic>
  <Term>VEGETATION</Term>
  <Variable_Level_1>EVERGREEN VEGETATION<
/Variable_Level_1>
</Science_Keywords>
```

UMM

```
ScienceKeywords: [
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "VEGETATION INDEX"
    VariableLevel2: "NORMALIZED DIFFERENCE
VEGETATION INDEX (NDVI)"
    DetailedVariable: "0.9 DENSITY"
  },
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "EVERGREEN VEGETATION"
  }
],
```

ECHO 10

Science Keywords are required. An unlimited amount of Science Keywords may be provided (Cardinality: 1..*)

UMM-C Element	ECHO 10 Path	Type	Constraints	Required in ECHO10?	Cardinality	Notes
ScienceKeywords /Category	/Collection/ScienceKeywords/ScienceKeyword/CategoryKeyword	String	KMS controlled	Yes	1	The category keyword will always be "EARTH SCIENCE"
ScienceKeywords /Topic	/Collection/ScienceKeywords/ScienceKeyword/TopicKeyword	String	KMS controlled	Yes	1	
ScienceKeywords /Term	/Collection/ScienceKeywords/ScienceKeyword/TermKeyword	String	KMS controlled	Yes	1	
ScienceKeywords /VariableLevel1	/Collection/ScienceKeywords/ScienceKeyword /VariableLevel1Keyword/Value	String	KMS controlled	No	0..1	
ScienceKeywords /VariableLevel2	/Collection/ScienceKeywords/ScienceKeyword /VariableLevel1Keyword/VariableLevel2Keyword/Value	String	KMS controlled	No	0..1	
ScienceKeywords /VariableLevel3	/Collection/ScienceKeywords/ScienceKeyword /VariableLevel1Keyword/VariableLevel2Keyword /VariableLevel3Keyword/Value	String	KMS controlled	No	0..1	
ScienceKeywords /DetailedVariable	/Collection/ScienceKeywords/ScienceKeyword /DetailedVariableKeyword	String	Uncontrolled (Free-Text)	No	0..1	

Example Mapping

ECHO 10

```
<ScienceKeyword>
  <CategoryKeyword>EARTH SCIENCE</CategoryKeyword>
  <TopicKeyword>BIOSPHERE</TopicKeyword>
  <TermKeyword>VEGETATION</TermKeyword>
  <VariableLevel1Keyword>
    <Value>VEGETATION INDEX</Value>
      <VariableLevel2Keyword>
        <Value>NORMALIZED DIFFERENCE VEGETATION
INDEX (NDVI)</Value>
      </VariableLevel2Keyword>
  </VariableLevel1Keyword>
  <DetailedVariableKeyword>0.9 DENSITY<
/DetailedVariableKeyword>
</ScienceKeyword>
<ScienceKeyword>
  <CategoryKeyword>EARTH SCIENCE</CategoryKeyword>
  <TopicKeyword>BIOSPHERE</TopicKeyword>
  <TermKeyword>VEGETATION</TermKeyword>
  <VariableLevel1Keyword>
    <Value>EVERGREEN VEGETATION</Value>
  </VariableLevel1Keyword>
</ScienceKeyword>
```

UMM

```
ScienceKeywords: [
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "VEGETATION INDEX"
    VariableLevel2: "NORMALIZED DIFFERENCE
VEGETATION INDEX (NDVI)"
    DetailedVariable: "0.9 DENSITY"
  },
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "EVERGREEN VEGETATION"
  }
],
```

ISO 19115-2 MENDS

Science Keywords are required. An unlimited amount of Science Keywords may be provided (Cardinality: 1..*)

UMM-C Element	ISO 19115-2 MENDS Path	Type	Notes

ScienceKeywords /Category ScienceKeywords /Topic ScienceKeywords /Term ScienceKeywords /VariableLevel1 ScienceKeywords /VariableLevel2 ScienceKeywords /VariableLevel3 ScienceKeywords /DetailedVariable	/gmi:MI_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords /gmd:MD_Keywords/ gmd:keyword/gco:CharacterString (list each value of the keyword hierarchy delimited by >)	String	KMS controlled. This is where the entire keyword hierarchy should be listed. Each keyword in the hierarchy must be separated by ">". If any keyword is missing and there exists a keyword later in the hierarchy (such as DetailedLocation), use NONE to fill in the values in between. The CMR will not translate the NONE values they are only used to place each keyword in its correct space in the hierarchy.
	/gmi:MI_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords /gmd:MD_Keywords/ gmd:keyword/gmd:type /MD_KeywordTypeCode[@codeListValue="theme"]	Codelist	codeList=https://cdn.earthdata.nasa.gov/iso/resources/Codelist/gmxCodelists.xml#MD_KeywordTypeCode Select the value "theme" from the codelist. This codelist value does not directly map to a UMM element; choosing "theme" indicates to CMR that the Science Keywords should be mapped.

Example Mapping

ISO 19115-2 MENDS

```
<gmi:MI_Metadata>
  ...
  <gmd:identificationInfo>
    <gmd:MD_DataIdentification>
      <gmd:descriptiveKeywords>
        <gmd:MD_Keywords>
          <gmd:keyword>
            <gco:CharacterString>EARTH SCIENCE&gt;
BIOSPHERE&gt;VEGETATION&gt;VEGETATION INDEX&gt;
NORMALIZED DIFFERENCE VEGETATION INDEX (NDVI)&gt;
NONE&gt;0.9 DENSITY</CharacterString>
          </gmd:keyword>
          <gmd:keyword>
            <gco:CharacterString>EARTH SCIENCE&gt;
BIOSPHERE&gt;VEGETATION&gt;EVERGREEN VEGETATION<
/CharacterString>
          </gmd:keyword>
          <gmd:type>
            <gmd:MD_KeywordTypeCode codeList="
https://cdn.earthdata.nasa.gov/iso/resources/Codelist
/gmxCodelists.xml#MD_KeywordTypeCode" codeListValue="
theme">theme</gmd:MD_KeywordTypeCode>
          </gmd:type>
        </gmd:MD_Keywords>
      </gmd:descriptiveKeywords>
      ...
```

UMM

```
ScienceKeywords: [
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "VEGETATION INDEX"
    VariableLevel2: "NORMALIZED DIFFERENCE
VEGETATION INDEX (NDVI)"
    DetailedVariable: "0.9 DENSITY"
  },
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "EVERGREEN VEGETATION"
  }
],
```

ISO 19115-2 SMAP

Science Keywords are required. An unlimited amount of Science Keywords may be provided (Cardinality: 1..*)

UMM-C Element	ISO 19115-2 SMAP Path	Type	Notes
ScienceKeywords /Category ScienceKeywords /Topic ScienceKeywords /Term ScienceKeywords /VariableLevel1 ScienceKeywords /VariableLevel2 ScienceKeywords /VariableLevel3 ScienceKeywords /DetailedVariable	/gmd:DS_Series/gmd:seriesMetadata/gmi:MI_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords/ gmd:MD_Keywords/gmd:keyword/gco:CharacterString (list each value of the keyword hierarchy delimited by >)	String	KMS controlled. This is where the entire keyword hierarchy should be listed. Each keyword in the hierarchy must be separated by ">". If any keyword is missing and there exists a keyword later in the hierarchy (such as DetailedLocation), use NONE to fill in the values in between. The CMR will not translate the NONE values they are only used to place each keyword in its correct space in the hierarchy.
	/gmi:MI_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:descriptiveKeywords/gmd:MD_Keywords/gmd:keyword/ gmd:type /MD_KeywordTypeCode[@codeListValue="theme"]	Codelist	codeList=https://cdn.earthdata.nasa.gov/iso/resources/Codelist/gmxCodelists.xml#MD_KeywordTypeCode Select the value "theme" from the codelist. This codelist value does not directly map to a UMM element; choosing "theme" indicates to CMR that the Science Keywords should be mapped.

Example Mapping

ISO 19115-2 SMAP

```
<gmd:DS_Series>
 <gmd:seriesMetadata>
  <gmi:MI_Metadata>
    ...
    <gmd:identificationInfo>
      <gmd:MD_DataIdentification>
        <gmd:descriptiveKeywords>
          <gmd:MD_Keywords>
            <gmd:keyword>
              <gco:CharacterString>EARTH SCIENCE&gt;
BIOSPHERE&gt;VEGETATION&gt;VEGETATION INDEX&gt;
NORMALIZED DIFFERENCE VEGETATION INDEX (NDVI)&gt;
NONE&gt;0.9 DENSITY</CharacterString>
            </gmd:keyword>
            <gmd:keyword>
              <gco:CharacterString>EARTH SCIENCE&gt;
BIOSPHERE&gt;VEGETATION&gt;EVERGREEN VEGETATION<
/CharacterString>
            </gmd:keyword>
            <gmd:type>
              <gmd:MD_KeywordTypeCode codeList="
https://cdn.earthdata.nasa.gov/iso/resources/Codelist
/gmxCodelists.xml#MD_KeywordTypeCode" codeListValue="
theme">theme</gmd:MD_KeywordTypeCode>
            </gmd:type>
          </gmd:MD_Keywords>
        </gmd:descriptiveKeywords>
        ...
```

UMM

```
ScienceKeywords: [
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "VEGETATION INDEX"
    VariableLevel2: "NORMALIZED DIFFERENCE
VEGETATION INDEX (NDVI)"
    DetailedVariable: "0.9 DENSITY"
  },
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "EVERGREEN VEGETATION"
  }
],
```

UMM Migration

None

Future Mappings

ISO 19115-1

Science Keywords are required. An unlimited amount of Science Keywords may be provided (Cardinality: 1..*)

UMM-C Element	ISO 19115-1 Path	Type	Notes

ScienceKeywords/Category ScienceKeywords/Topic ScienceKeywords/Term ScienceKeywords/VariableLevel1 ScienceKeywords/VariableLevel2 ScienceKeywords/VariableLevel3 ScienceKeywords/DetailedVariable	/mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:descriptiveKeywords/ mri:MD_Keywords/mri:keyword/gco:CharacterString (list each value of the keyword hierarchy delimited by >)	String	KMS controlled. This is where the entire keyword hierarchy should be listed. Each keyword in the hierarchy must be separated by ">". If any keyword is missing and there exists a keyword later in the hierarchy (such as DetailedLocation), use NONE to fill in the values in between. The CMR will not translate the NONE values they are only used to place each keyword in its correct space in the hierarchy.
	/mdb:MD_Metadata/mdb:identificationInfo/mri:MD_DataIdentification/mri:descriptiveKeywords/ mri:MD_Keywords/mri:keyword/mri:type/mri:MD_KeywordTypeCode[@codeListValue="theme"]	Codelist	codeList=http://standards.iso.org/iso/19115/resources/Codelist/cat/codelists.xml#MD_KeywordTypeCode Select the value "theme" from the codelist. This codelist value does not directly map to a UMM element; choosing "theme" indicates to CMR that the Science Keywords should be mapped.

Example Mapping

ISO 19115-1

```
<mdb:MD_Metadata>
  ...
    <mdb:identificationInfo>
        <mri:MD_DataIdentification>

            <mri:descriptiveKeywords>
                <mri:MD_Keywords>
                    <mri:keyword>
                        <gco:CharacterString>EARTH
SCIENCE&gt;BIOSPHERE&gt;VEGETATION&gt;VEGETATION
INDEX&gt;NORMALIZED DIFFERENCE VEGETATION INDEX
(NDVI)&gt;NONE&gt;0.9 DENSITY</gco:CharacterString>
                    </mri:keyword>
                    <mri:keyword>
                        <gco:CharacterString>EARTH
SCIENCE&gt;BIOSPHERE&gt;VEGETATION&gt;EVERGREEN
VEGETATION</gco:CharacterString>
                    </mri:keyword>
                    <mri:type>
                        <mri:MD_KeywordTypeCode
                            codeList="
http://standards.iso.org/iso/19115/resources/Codelist
/cat/codelists.xml#MD_KeywordTypeCode"
                            codeListValue="
MD_KeywordTypeCode_theme">theme</mri:
MD_KeywordTypeCode>
                    </mri:type>
                </mri:MD_Keywords>
            </mri:descriptiveKeywords>
            ...
```

UMM

```
ScienceKeywords: [
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "VEGETATION INDEX"
    VariableLevel2: "NORMALIZED DIFFERENCE
VEGETATION INDEX (NDVI)"
    DetailedVariable: "0.9 DENSITY"
  },
  {
    Category: "EARTH SCIENCE",
    Topic: "BIOSPHERE",
    Term: "VEGETATION",
    VariableLevel1: "EVERGREEN VEGETATION"
  }
],
```

History

UMM Versioning

Version	Date	What Changed
1.15.5	12/3/2020	No changes were made for Science Keywords during the transition from version 1.15.4 to 1.15.5
1.15.4	9/18/2020	No changes were made for Science Keywords during the transition from version 1.15.3 to 1.15.4
1.15.3	7/1/2020	No changes were made for Science Keywords during the transition from version 1.15.2 to 1.15.3
1.15.2	5/20/2020	No changes were made for Science Keywords during the transition from version 1.15.1 to 1.15.2
1.15.1	3/25/2020	No changes were made for Science Keywords during the transition from version 1.15.0 to 1.15.1
1.15.0	2/26/2020	No changes were made for Science Keywords during the transition from version 1.14.0 to 1.15.0
1.14.0	10/21/2019	No changes were made for Science Keywords during the transition from version 1.13.0 to 1.14.0
1.13.0	04/11/2019	No changes were made for Science Keywords during the transition from version 1.12.0 to 1.13.0
1.12.0	01/22/2019	No changes were made for Science Keywords during the transition from version 1.11.0 to 1.12.0.
1.11.0	11/28/2018	No changes were made for Science Keywords during the transition from version 1.10.0 to 1.11.0.
1.10.0	05/02/2018	No changes were made for Science Keywords during the transition from version 1.9.0 to 1.10.0.

ARC Documentation

Version	Date	What Changed	Author
1.0	05/04/2018	Recommendations/priority matrix transferred from internal ARC documentation to wiki space	Jeanne' le Roux