

NASA's Global Change Master Directory (GCMD)

Directory Interchange Format (DIF) Writer's Guide

This guide was written by the Global Change Master Directory staff to assist with the creation of data set descriptions. Any material in this pdf document may be superseded by online published content [<http://gcmd.nasa.gov/User/difguide/>]. If you have any questions please contact GCMD (gcmduso@gcmd.nasa.gov).

This document should be cited as:

Directory Interchange Format (DIF) Writer's Guide, 2008.

Global Change Master Directory. National Aeronautics and Space Administration.

Contents

	What is a DIF?	3
Required	Entry ID	5
	Entry Title	6
	Parameters (Science Keywords)	7
	ISO Topic Category	10
	Data Center	12
	Summary	15
	Metadata Name	16
	Metadata Version	18
Highly Recommended	Data Set Citation	19
	Personnel	22
	Instrument	26
	Platform	28
	Temporal Coverage	30
	Paleo-Temporal Coverage	32
	Spatial Coverage	35
	Location	38
	Data Resolution	41
	Project	43
	Quality	45
	Access Constraints	46
	Use Constraints	47
	Distribution	48
	Data Set Language	50
	Data Set Progress	51
	Related URL	52
	DIF Revision History	54
Recommended	Keyword (Ancillary Keyword)	55
	Originating Center	56
	Multimedia Sample	57
	References (Publications)	59
	Parent DIF	60
	IDN Node	62
	DIF Creation Date	63
	Last DIF Revision Date	64
	Future DIF Review Date	65
	Privacy Status	66

What is a DIF?

A Short History of the Directory Interchange Format (DIF)

- by Lola Olsen

Before metadata became a ubiquitous buzzword, a descriptive and standardized format for exchanging information about scientific data sets was conceived and implemented. The Directory Interchange Format, the DIF, was the product of an Earth Science and Applications Data Systems Workshop (ESADS) held February 24-26, 1987 on catalog interoperability (CI). The workshop recommended that a "...first step towards data system interoperability, Catalog Interoperability (CI), the ability to find information about data held at other sites...", be made.

In the summer of 1987, the Catalog Interoperability Working Group (consisting of several U.S. Federal and international agencies) defined the type of information and level of detail that would be contained by the DIF. The DIF structure was "frozen" on September 18, 1987, and the population of NASA's Master Directory (NMD) prototype commenced. By December 1987, over 100 DIF entries were available in the prototype NMD database. After several demonstrations, workshops, and feedback from the scientific community, the Directory Interchange Format (DIF) was formally approved and adopted by a CI science advisory group at a CI workshop in 1988.

In 1989, the Committee on Earth Observation Satellites (CEOS) Data Working Group (DWG) began attending the CI Workshop meetings and provided valuable feedback on the DIF structure. The CEOS International Directory Network (IDN) was soon established under the auspices of the CEOS Working Group on Data, through the Catalog Subgroup, to foster the exchange of information among international agencies. The first release of the IDN was named the Prototype International Directory (PID) in 1990. [Actual DIF exchange procedures were agreed on by February 1991.]

In 1990, the Interagency Working Group on Data Management for Global Change (IWGDMGC) adopted the directory as a prototype to facilitate global change research - in response to the challenge by the Earth System Science Committee (ESSC). Thereafter, the NMD was renamed the Global Change Master Directory (GCMD) for its Earth sciences applications.

The DIF has enjoyed over 16 years of success. The DIF structure has been flexible enough to evolve with growing metadata requirements, especially for the geospatial disciplines. In the U.S., the Office of Management and Budget's (OMB's) Circular A-16 for the improved coordination of spatial data among federal agencies led to the establishment of the Federal Geographic Data Committee (FGDC) and the FGDC Clearinghouse. The GCMD serves as NASA's FGDC Clearinghouse node for geospatial metadata. Elements of the Content Standard for Digital Geospatial Metadata (CSDGM) were incorporated in the DIF in 1994.

In the late 1990s, the geospatial community began work towards the development of an international standard for geospatial metadata. The International Organisation for Standardisation (ISO) Technical Committee's [(TC)211] Metadata Standard 19115 (previously known as 15046-15) (see: <http://www.isotc211.org/>) sought to provide "a consistent suite of geographic information schemata that allows geographic information to be integrated with information technology. The goal of this work item is to produce a schema for geographic information metadata."

The ISO 19115/TC211 geospatial metadata standard was adopted June 2004. Required elements and appropriate modifications were approved by the CEOS IDN Interoperability group and incorporated into the DIF to achieve full ISO compatibility.

The DIF does not compete with other metadata standards. It is simply the "container" for the metadata elements that are maintained in the IDN database, where validation for mandatory fields, keywords, personnel, etc. takes place.

The DIF is used to create directory entries which describe a group of data. A DIF consists of a collection of fields which detail specific information about the data. Eight fields are required in the DIF; the others expand upon and clarify the information. Some of the fields are text fields, others require the use of controlled keywords (sometimes known as "valids").

The DIF allows users of data to understand the contents of a data set and contains those fields which are necessary for users to decide whether a particular data set would be useful for their needs.

The DIF Guide document provides information about each field of the DIF, including its syntax, specifications, recommendations, and examples. Several example DIFs are also provided.

Entry Identifier

Definition

The **<Entry_ID>** is the unique identifier of the metadata record. The **<Entry_ID>** is determined by the metadata author and may be identical to identifiers used by the data provider's data center or organization. For example, the National Snow and Ice Data Center (NSIDC) DAAC identifies their metadata records as *NSIDC-xxxx*, where *xxxx* is a numerical designator. Also, the identifier is case insensitive meaning *nsidc-xxxx* and *NSIDC-xxx* refer to the same metadata record.

Syntax

<Entry_ID> text entry identifier **</Entry_ID>**

- The **<Entry_ID>** consists of 1 to 80 alphanumeric characters of the UTF-8 set, including underbar (_), hyphen (-) and period (.)
- The following characters are not allowed in the **<Entry_ID>**:
 - Backward Slash (\)
 - Forward Slash (/)
 - Colon (:)
 - Spaces ()

Usage

- The **<Entry_ID>** is required and is not repeatable.

Examples

<Entry_ID>OAR0005**</Entry_ID>**

<Entry_ID>JPL_DAAC_SSMI_WINDS**</Entry_ID>**

The following would **not** be allowed:

<Entry_ID>NSIDC23/5**</Entry_ID>**. The forward slash (/) is not an allowed character.

Entry Title

Definition

The **<Entry_Title>** is the title of the data set described by the metadata.

- **<Entry_Title>** should be descriptive enough so that when a user is presented with a list of titles the general content of the data set can be determined. For example, **<Entry_Title>Aerosols</Entry_Title>** would not be an adequate data set title as it does not provide enough descriptive information to guide the user.
- In order to make titles descriptive, important elements about the data may be included in the **<Entry_Title>**, i.e., parameters measured, geographic location, instrument, investigator, project, temporal coverage. For example, **<Entry_Title>Aerosol characterization and snow chemistry at Terra Nova Bay 2001-2003</Entry_Title>** provides an adequate amount of information to guide the user.

Syntax

<Entry_Title>Title of the data set**</Entry_Title>**

- 1 to 220 characters permitted from the entire UTF-8 character set.
- The **<Entry_Title>** should follow standard mixed (i.e. upper and lower) case.

Usage

The **<Entry_Title>** field is required and is not repeatable.

Examples

<Entry_Title>Hourly Tidal Observations from the Coast of China, 1942-1986**</Entry_Title>**

<Entry_Title>The San Joaquin Valley Air Quality Study (SJVAQS) Atmospheric Signatures, Predictions, and Experiments (AUSPEX) Database**</Entry_Title>**

<Entry_Title>The Next Generation Weather Radar (NEXRAD) Level II, III, and IV Data Products**</Entry_Title>**

Parameters (Science Keywords)

Definition

The **<Parameters>** field allows for the specification of Earth science keywords that are representative of the data set being described. These keywords are important for the precise search and retrieval of information from the GCMD. The author must select these keywords from the controlled set of science keywords. The **<Parameters>** field consists of a 7-level hierarchical classification of science keywords, defined as follows:

- **<Category>** is the highest keyword category. The default is "*EARTH SCIENCE*".
- **<Topic>** is the next highest level in the keyword hierarchy under **<Category>**.
There are 14 Topics:
 - Agriculture
 - Atmosphere
 - Biosphere
 - Biological Classification
 - Climate Indicators
 - Cryosphere
 - Human Dimensions
 - Land Surface
 - Oceans
 - Paleoclimate
 - Solid Earth
 - Spectral/Engineering
 - Sun-Earth Interactions
 - Terrestrial Hydrosphere
- **<Term>** is next level in the hierarchy under **<Topic>**. There can be several controlled **<Term>** keywords for each **<Topic>**. For example,
<Topic>Atmosphere</Topic> <Term>Aerosols </Term>.
- **<Variable_Level_1>** is the next level in the hierarchy under **<Term>**. There can be several controlled **<Variable_Level_1>** keywords for each **<Term>**. For example, <Topic>Atmosphere</Topic> <Term>Aerosols </Term>
<Variable_Level_1>Aerosol Optical Depth</Variable_Level_1>.
- **<Variable_Level_2>** is the next level in the hierarchy under **<Variable_Level_1>**. There can be several controlled **<Variable_Level_2>** keywords for each **<Variable_Level_1>**. For example, <Topic>Biological Classification</Topic>

<Term>Animals/Invertebrates</Term> <Variable_Level_1>Arthropods
</Variable_Level_1> <Variable_Level_2>Crustaceans</Variable_Level_2>.

- **<Variable_Level_3>** is the next level in the hierarchy under **<Variable_Level_2>**. There can be several controlled **<Variable_Level_3>** keywords for each **<Variable_Level_2>**. For example, <Topic>Biological Classification</Topic>
<Term> Animals/Invertebrates </Term>
<Variable_Level_1>Arthropods</Variable_Level_1>
<Variable_Level_2>Crustaceans</Variable_Level_2>
<Variable_Level_3>Decapods </Variable_Level_3>
- **<Detailed_Variable>** is an uncontrolled free text field that allows the metadata author to specify any keywords to more exactly describe the measurement represented by the data. For example, <Topic>Biological Classification</Topic>
<Term> Animals/Invertebrates </Term>
<Variable_Level_1>Arthropods</Variable_Level_1>
<Variable_Level_2>Crustaceans</Variable_Level_2>
<Variable_Level_3>Decapods </Variable_Level_3> <Detailed_Variable>Lobsters
</Detailed_Variable>

Syntax

<Parameters>

<Category>EARTH SCIENCE </Category> (*by default*)

<Topic>Topic keyword </Topic>

<Term> Term keyword </Term>

<Variable_Level_1> Variable Level 1 keyword </Variable_Level_1>

<Variable_Level_2> Variable Level 2 keyword </Variable_Level_2>

<Variable_Level_3> Variable Level 3 keyword </Variable_Level_3>

<Detailed_Variable> text </Detailed_Variable>

</Parameters>

- **<Detailed_Variable>** is a free-text field, which may contain up to 80 characters.
- **<Category>**, **<Topic>**, **<Term>**, **<Variable_Level_1>**, **<Variable_Level_2>**, and **<Variable_Level_3>** must be selected from the set of controlled Earth science keywords.

Usage

The `<Parameters>` field is required and repeatable.

The following fields within `<Parameters>` are required and may not be repeated within the `<Parameters>` field:

`<Category>`

`<Topic>`

`<Term>`

The following fields within `<Parameters>` are optional and may not be repeated within the `<Parameters>` field:

`<Variable_Level_1>`

`<Variable_Level_2>`

`<Variable_Level_3>`

`<Detailed_Variable>`

Multiple **`<Parameters>`** can be specified for each metadata record. All applicable **`<Topic>`**, **`<Term>`**, **`<Variable_Level_1>`**, **`<Variable_Level_2>`**, and **`<Variable_Level_3>`** science keywords that describe the data set should be selected.

Examples

Specification of `<Parameters>` without a `<Variable>` keyword.

```
<Parameters>
<Category>Earth Science</Category>
<Topic>Atmosphere</Topic>
<Term>Precipitation</Term>
</Parameters>
```

Specification of `<Parameters>` with a `<Variable_Level_1>` keyword.

```
<Parameters>
<Category>Earth Science</Category>
<Topic>Climate Indicators</Topic>
<Term>Teleconnections</Term>
<Variable_Level_1>North Atlantic Oscillation</Variable_Level_1>
</Parameters>
```

ISO Topic Category

Definition

The **<ISO_Topic_Category>** field is used to identify the keywords in the ISO 19115 - Geographic Information Metadata (<http://www.isotc211.org/>) Topic Category Code List. It is a high-level geographic data thematic classification to assist in the grouping and search of available geographic data sets. The **<ISO_TOPIC_Category>** keywords are as follows:

- Farming
- Biota
- Boundaries
- Climatology/Meteorology/Atmosphere
- Economy
- Elevation
- Environment
- Geoscientific Information
- Health
- Imagery/Base Maps/Earth Cover
- Intelligence/Military
- Inland Waters
- Location
- Oceans
- Planning Cadastre
- Society
- Structure
- Transportation
- Utilities/Communications

For definitions of these keywords, see:

http://gcmd.nasa.gov/User/difguide/iso_topics.html

Syntax

<ISO_Topic_Category>ISO Topic Category keywords**</ ISO_Topic_Category>**

Usage

The **<ISO_Topic_Category>** field is required and may be repeated.

The **< ISO_Topic_Category >** must be selected from the ISO Topic Category list of keywords (see above) from the ISO 19115 metadata standard (MD_TopicCategory Code).

Example

<ISO_Topic_Category>Farming</ISO_Topic_Category>

Data Center

Definition

The **<Data Center>** is the data center, organization, or institution responsible for distributing the data.

- **<Data_Center_Name>** consists of the data center **<Short_Name>** and **<Long_Name>**, which is the name of the data center that distributes the data.
- **<Data_Center_URL>** is the URL of the data center.
- **<Data_Set_ID>** is a data set identifier assigned by the data center (may or may not be the same as the **<Entry_ID>**).
- **<Personnel>** Contact information for the data including name, email, phone, FAX, and address information. See the [<Personnel>](#) entry for a complete description.

Syntax

```
<Data_Center>
<Data_Center_Name>
<Short_Name>Data Center short name from the controlled list,
    up to 160 characters
</Short_Name>
<Long_Name>Data Center long name from controlled list,
    up to 240 characters
</Long_Name>
</Data_Center_Name>
<Data_Center_URL>URL up to 600 characters</Data_Center_URL>
<Data_Set_ID>Text up to 80 characters</Data_Set_ID>
<Personnel>
<Role>DATA CENTER CONTACT</Role>
<First_Name>Text up to 80 characters</First_Name>
<Middle_Name>Text up to 80 characters</Middle_Name>
<Last_Name>Text up to 80 characters</Last_Name>
<Email>Text up to 80 characters, of the form janedoe@domain</Email>
<Phone>Text up to 80 characters of the form "xxx-xxx-xxxx".
    Outside of the US and Canada, the country code should
    be added (+xx-xxx-xxxx)
</Phone>
<Fax>Text up to 80 characters (see syntax for </Fax>
<Contact_Address>
<Address>Text up to 80 characters </Address>
<City>Text up to 80 characters </City>
```

<Province_or_State>Text up to 80 characters</Province_or_State>
<Postal_Code>Text up to 80 characters</Postal_Code>
<Country>Text up to 80 characters </Country>
</Contact_Address>
</Personnel>
</Data_Center>

- Characters may be selected from the UTF-8 character set.

Usage

The **<Data_Center>** field is required and can be repeated.

<Data_Center_Name> may only be used once per **<Data Center>**. It consists of the data center **<Short_Name>** and **<Long_Name>**. If the **<Short_Name>** does not appear in the list of controlled data center keywords, a new data center **<Short_Name>** may be submitted.

<Data_Center_URL> is a required subfield.

<Data_Set_ID> is an optional subfield and may be repeated within the **<Data Center>** field as often as necessary.

<Personnel> is a required subfield and may be repeated within the **<Data Center>** field as often as necessary. See the **<Personnel>** entry for usage.

<Role> is required and is not repeatable. The only allowable **<Role>** within **<Data_Center>** is *"Data Center Contact"*.

<Email>, **<Phone>** and **<FAX>** are optional and repeatable. See the **<Personnel>** entry for usage.

<Address> may be repeated as often as necessary. See the **<Personnel>** entry for usage.

Example

```
<Data_Center>
  <Data_Center_Name>
    <Short_Name>LP DAAC</Short_Name>
    <Long_Name>Land Processes Distributed Active
      Archive Center</Long_Name>
  </Data_Center_Name>
  <Role>DATA CENTER CONTACT</Role>
  <Phone>(Toll Free) (866) 573-3222</Phone>
  <Phone>(605) 594-6116</Phone>
  <Fax>(605) 594-6963</Fax>
```

<Address>LP DAAC User Services</Address>
<Address>U.S. Geological Survey (USGS)</Address>
<Address>Center for Earth Resources Observation
and Science(EROS)</Address>
<Address>47914 252nd Street</Address>
<City>Sioux Falls</City>
<Province_or_State>SD</Province_or_State>
<Postal_Code>57198-0001</Postal_Code>
<Country>USA</Country>
</Contact_Address>
</Personnel>
</Data_Center>

Summary

Definition

The **<Summary>** is a brief description of the data set that allows potential users to determine if the data set is useful for their needs. The following are suggestions to populate the **<Summary>** field:

- 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.
- The **<Summary>** may contain tabular information.
- URLs will be automatically hyperlinked.
- Where applicable, the **<Summary>** should include brief statements of 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).
 - 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.
 - Purpose and/or intended use of data.
 - Other pertinent information.

Syntax

<Summary>text</Summary>

- Multiple lines of text are allowed for the **<Summary>** field and may contain characters from the UTF-8 character set.

Usage

The **<Summary>** field is required and may not be repeated.

Examples

<Summary>Satellite Coincidence Subsetted Intermediate Data (CSI): A subset of 1B01, calibrated VIRS (0.63, 1.6, 3.75, 10.8, and 12 um) radiances at 2.2 km resolution over a 720 km swath from a TRMM pass over the SCSMEX Southern Site.</Summary>

<Summary>Global monthly AMSR-E-derived Soil Moisture.

The dataset contains global monthly-mean soil moisture statistics (average values and standard deviation) for 1 by 1 degree grid cells. The source for the data is AMSR-E daily estimates of soil moisture (Product name: AMSR_E_L3_DailyLand). The dataset covers the time period starting July 2002.

Data files contain 180 lines with 360 pixels per line. The projection is latitude-longitude with the upper left corner of the first pixel of the first line positioned at 90°N 180°W.

</Summary>

Metadata Name

Definition

The ISO 19115 **<Metadata_Name>** field is used to identify the current DIF standard name. This field is automatically populated when using the GCMD metadata authoring tool.

Syntax

<Metadata_Name>CEOS IDN DIF**</Metadata_Name>** (*by default*)

- The **<Metadata_Name>** field may contain 1 to 80 characters from the UTF-8 character set.

Usage

The **<Metadata_Name>** field is required and may not be repeated.

This field is automatically populated (with **CEOS IDN DIF**) when using the GCMD metadata authoring tool.

Example

<Metadata_Name>CEOS IDN DIF**</Metadata_Name>**

Metadata Version

Definition

The **<Metadata_Version>** field is used to identify the current DIF metadata standard. This field is automatically populated when using the GCMD metadata authoring tools.

Syntax

`<Metadata_Version>9.7</Metadata_Version>` (*by default*)

- The **<Metadata_Version>** field may contain 1 to 80 characters from the UTF8 character set.

Usage

The **<Metadata_Version>** field is required and may not be repeated.
This field is automatically populated when using the GCMD metadata authoring tool.

Example

`<Metadata_Version>9.7</Metadata_Version>`

Data Set Citation

Definition

The **<Data_Set_Citation>** field allows the author to properly cite the data set producer. This field has 2 functions:

1. To indicate how this data set should be cited in the professional scientific literature, and
2. If this data set is a compilation of other data sets, to document and credit the data sets that were used in producing this compilation.

This field is not to be used to list bibliographic references of scientific research articles arising from the data set. This field provides a citation for the data set itself, not articles related to the research results. To list references related to the research results, use the **<Reference>** field. **<Data_Set_Citation>** consists of:

- **<Dataset_Creator>** The name of the organization(s) or individual(s) with primary intellectual responsibility for the data set's development.
- **<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. For example **<Dataset_Series_Name>IGBP PAGES/WDC CONTRIBUTION SERIES</Dataset_Series_Name>**
- **<Dataset_Release_Date>** The date when the data set was made available for release.
- **<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. For example, **<Version>Version 1.2</Version>**
- **<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.
- **<Online_Resource>** The URL of the online resource containing the data set.

Syntax

<Data_Set_Citation>
<Dataset_Creator>Text, names separated by commas up to 500 characters</Dataset_Creator>
<Dataset_Title>Text up to 220 characters</Dataset_Title>
<Dataset_Series_Name>Text up to 220 characters</Dataset_Series_Name>
<Dataset_Release_Date>Text up to 31 characters, yyyy-mm-dd format suggested</Dataset_Release_Date>
<Dataset_Release_Place>Text up to 80 characters</Dataset_Release_Place>
<Dataset_Publisher>Text up to 500 characters</Dataset_Publisher>
<Version> Number and/or text</Version>
<Issue_Identification>Number and/or text up to 80 characters</Issue_Identification>
<Data_Presentation_Form>Text up to 80 characters</Data_Presentation_Form>
<Other_Citation_Details>Text up to 160 characters</Other_Citation_Details>
<Online_Resource>URL up to 600 characters</Online_Resource>
</Data_Set_Citation>

- The **<Dataset_Creator>** and the **<Dataset_Publisher>** fields may contain up to 500 characters. The **<Dataset_Title>** field may contain up to 220 characters. The **<Dataset_Series_Name>** and **<Other_Citation_Details>** fields may contain up to 160 characters from the UTF-8 character set. The **<Online_Resource>** field may contain up to 600 characters. The **<Dataset_Release_Date>** field may contain up to 31 characters.
- All other fields may contain 1 to 80 characters from the UTF-8 character set.

Usage

The **<Data_Set_Citation>** is a highly recommended field and can be repeated. Individual subfields may not be repeated.

Keywords for **<Data_Presentation_Form>** may be selected from the Suggested Keywords List.

Example

<Data_Set_Citation>
<Dataset_Creator>Fred K. Miller, Jonathan C. Matti,
Howard J. Brown, Pamela M. Cossette</Dataset_Creator>
<Dataset_Title>Digital Geologic Map of the Butler Peak 7.5'
Quadrangle, San Bernardino County, California</Dataset_Title>

<Dataset_Series_Name>U.S. Geological Survey Open-File
Report</Dataset_Series_Name>
<Dataset_Release_Date>2000-01-30</Dataset_Release_Date>
<Dataset_Release_Place>Menlo Park, California</Dataset_Release_Place>
<Dataset_Publisher>U.S. Geological Survey</Dataset_Publisher>
<Version>1.0</Version>
<Data_Presentation_Form>vector digital data</Data_Presentation_Form>
<Online_Resource><http://geopubs.wr.usgs.gov/open-file/of00-145></Online_Resource>
</Data_Set_Citation>

Personnel

Definition

<Personnel> defines the point of contact for more information about the data set or the metadata.

- The contact personnel are defined by the **<Role>**, which include:
 - **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)
 - **Technical Contact**: The person who is knowledgeable about the technical content of the data (quality, processing methods, units, available software for further processing)
 - **DIF Author**: The person who is responsible for the content of the DIF. If the responsibility shifts from the original author to another person, the DIF Author field should be updated to the new responsible person.

- **<First_Name>**, **<Middle_Name>** and **<Last_Name>** are the first, middle and last name of the person or organization defined in the **<Personnel>** field. Initials may be used for the **<First_Name>** and **<Middle_Name>**. Organizational names may be substituted for personal names.
- **<Email>** is the email address of the personnel or organization. Note: Authors may use “@” in formatting email addresses, however the “@” symbol will not be displayed to the public to avoid potential spam attacks.
- **<Phone>** is the telephone number of the person or organization. Telephone extensions are allowed. If not in the U.S. or Canada, use the two-digit country code followed by the phone number. Otherwise, use the 10-digit phone number including area code. For example: **<Phone>+44 5555 555555</Phone>** or **<Phone>301-555-5555</Phone>**
- **<FAX>** is the FAX number of the person or organization. The same rules for **<Phone>** apply to FAX numbers.
- **<Contact Address>** contain the address information of the person or organization. It consists of:
 - **<Address>** is the organization name, department, mail stop, street

address, etc. of the person or organization.

- **<City>** is the city or town of the person or organization. **<Province or State>** is the province (particularly Canadian provinces, region or state (particularly in the United States)).
- **<Postal Code>** is the postal code of the person or organization.
- **<Country>** is the country of the person or organization.

Syntax

<Personnel>

<Role>Personnel Role Valid**</Role>**

<First_Name>Text up to 80 characters**</First_Name>**

<Middle_Name>Text up to 80 characters**</Middle_Name>**

<Last_Name>Text up to 80 characters**</Last_Name>**

<Email>Text up to 80 characters of the form

"janedoe@domain"**</Email>**

<Phone>Text up to 80 characters of the form "xxx-xxx-xxxx". Outside of the US and Canada, the country code should be added

(+xx-xxx-xxxx)**</Phone>**

<Fax>Text up to 80 characters (see syntax for **<Phone>**)**</Fax>**

<Contact_Address>

<Address>Text up to 80 characters**</Address>**

<City>Text up to 80 characters**</City>**

<Province_or_State>Text up to 80 characters**</Province_or_State>**

<Postal_Code>Text up to 80 characters**</Postal_Code>**

<Country>Text up to 80 characters**</Country>**

</Contact_Address>

</Personnel>

- Characters may be selected from the UTF-8 character set.

Usage

The **<Personnel>** field is highly recommended and may be repeated.

The **<Role>** field is required and must be selected from the Personnel Role keyword list (see definitions above). The **<Role>** may be repeated. A person may have more than one role, such as **Investigator** and **DIF Author**.

The **<First_Name>** and **<Middle_Name>** of the personnel are optional. If the **<Personnel>** field is populated, the **<Last_Name>** is required. Names may be

personal names or organizational names.

The **<Email>**, **<Phone>** and **<Fax>** fields are optional and may be repeated.

The **<Address>** field may contain multiple lines and can be repeated as often as necessary.

<City>, **<Province_Or_State>**, **<Postal_Code>** and **<Country>** are not repeatable.

<Phone> and **<Fax>** are optional and may be repeated.

<Country> is optional and may not be repeated. **<Country>** should always be included in the address.

Note: If personnel retire, particularly investigators, it is desirable to keep their full name, delete their **<Phone>**, **<Email>** and **<Fax>** number, or in the **<Address>** subfield, state that they are retired or deceased, keeping the name of their affiliation.

Examples

Example of personnel with a role of DIF Author

```
<Personnel>
  <Role>DIF AUTHOR</Role>
  <First_Name>MICHAEL</First_Name>
  <Last_Name>MORAHAN</Last_Name>
  <Email>Michael.P.Morahan.1 @xxx.xxx.gov</Email>
  <Phone>301-555-5555</Phone>
  <Fax>301-222-2222</Fax>
  <Contact_Address>
    <Address>Goddard Earth Sciences Data and Information Services
    Center</Address>
    <Address>Code 610.2</Address>
    <Address>NASA Goddard Space Flight Center</Address>
    <City>Greenbelt</City>
    <Province_or_State>MD</Province_or_State>
    <Postal_Code>20771</Postal_Code>
    <Country>USA</Country>
  </Contact_Address>
</Personnel>
```

Example of personnel with role of Investigator and from non-US country.

```
<Personnel>
  <Role>Investigator</Role>
  <First_Name>Phil</First_Name>
```


<Middle_Name>D.</Middle_Name>
<Last_Name>Jones</Last_Name>
<Email>p.jone@zzz.zz.uk</Email>
<Phone>+44 5555 555555</Phone>
<Fax>+44 2222 222222</Fax>
<Contact_Address>
 <Address>Climatic Research Unit</Address>
 <Address>University of East Anglia</Address>
 <City>Norwich</City>
 <Postal_Code>NR4 7TJ</Postal_Code>
 <Country>United Kingdom</Country>
</Contact_Address>
</Personnel>

Example of organizational personnel with a role of Technical Contact

<Personnel>
 <Role>Technical Contact</Role>
 <Last_Name>WDC/PALEOCLIMATOLOGY, BOULDER</Last_Name>
 <Email>paleo@xxxx.gov</Email>
 <Phone>555-555-5555</Phone>
 <Fax>222-222-2222</Fax>
 <Contact_Address>
 <Address>World Data Center for Paleoclimatology</Address>
 <Address>NOAA/NCDC Code E/GCx3</Address>
 <Address>325 Broadway</Address>
 <City>Boulder</City>
 <Province or State>CO</Province or State>
 <Postal_Code>80303</Postal_Code>
 <Country>USA</Country>
 </Contact_Address>
</Personnel>

Instrument (Sensor Name)

Definition

The Instrument or **<Sensor_Name>** is the name of the instrument used to acquire the data. There are 3 categories of instruments, with additional levels of hierarchical classifications (see the Writer's Guide for Ancillary Descriptions):

- Earth Remote Sensing Instruments
- In Situ/Laboratory Instruments
- Solar/Space Observing Instruments

When using the GCMD metadata authoring tools, **<Sensor_Name>** classifications are automatically populated by selecting the **<Short_Name>**. This field allows for the specification of keywords that are the names of the instrument(s) used to collect or measure the data. These keywords are important for the search and retrieval of information from the GCMD. The field **<Sensor_Name>** consists of the short name and the long name of the instrument. There is a 1:1 correspondence between the short name and the long name.

- **<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.

Syntax

```
<Sensor_Name>  
  <Short_Name>Instrument short name keyword</Short_Name>  
  <Long_Name>Instrument long name keyword</Long_Name>  
</Sensor_Name>
```

Usage

The **<Sensor_Name>** field is highly recommended and may be repeated as many times as necessary.

<Short_Name> 1 to 80 characters from the UTF-8 character set, except ">"

<Long_Name> 1 to 160 characters from the UTF-8 character set, except ">"

<Short_Name> and **<Long_Name>** must be selected from the controlled set of Instrument keywords. However, new instrument identifiers may only be added if the instrument does not already appear in the Instrument keyword List.

The **<Short_Name>** is usually a mnemonic or abbreviated version of the

<Long_Name>. When submitting a new **<Short_Name>**, always check the Instrument Keyword Lists to ensure mutual exclusivity among instrument short names. No two instruments may have the same **<Short_Name>**.

Examples

```
<Sensor_Name>  
  <Short_Name>MODIS</Short_Name>  
  <Long_Name>Moderate-Resolution Imaging Spectroradiometer</Long_Name>  
</Sensor_Name>
```

```
<Sensor_Name>  
  <Short_Name>ADCP</Short_Name>  
  <Long_Name>Acoustic Doppler Current Profiler</Long_Name>  
</Sensor_Name>
```

Platform (Source Name)

Definition

The Platform or **<Source_Name>** is the name of the platform used to acquire the data, There are 11 categories of platforms, with additional levels of hierarchical classifications (see the Writer's Guide for Ancillary Descriptions):

- Aircraft
- Balloons/Rockets
- Earth Observation Satellites
- In Situ Land-based Platforms
- In Situ Ocean-based Platforms
- Interplanetary Spacecraft
- Maps/Charts/Photographs
- Models
- Navigation Platforms
- Solar/Space Observation Platforms
- Space Stations/Manned Spacecraft

When using the GCMD metadata authoring tools, **<Source_Name>** classifications are automatically populated by selecting the **<Short_Name>**. This field allows for the specification of keywords that are representative of the platform(s) or source used to collect the data. These keywords are important for the search and retrieval of information from the GCMD. The field **<Source_Name>** consists of the short name and the long name of the platform (source). There is a 1:1 correspondence between the short name and the long name.

- **<Short_Name>** is the abbreviated name of the platform (source) used to acquire the data.
- **<Long_Name>** is the full name of the platform (source) used to acquire the data.

Syntax

```
<Source_Name>  
  <Short_Name>Platform Short Name keyword</Short_Name>  
  <Long_Name>Platform Long Name keyword</Long_Name>  
</Source_Name>
```

Usage

The **<Source_Name>** field is highly recommended and may be repeated as many times as necessary.

<Short_Name> 1 to 80 characters from the UTF-8 character set, except ">"

<Long_Name> 1 to 160 characters from the UTF-8 character set, except ">"

<Short_Name> and **<Long_Name>** must be selected from the controlled set of Platform keywords. However, new platform identifiers may only be added if the platform does not already appear in the Platform Keyword List.

The **<Short_Name>** is usually a mnemonic or abbreviated version of the **<Long_Name>**. When submitting a new **<Short_Name>**, always check the Platform Keyword List to ensure mutual exclusivity among platform short names. No two platforms may have the same **<Short_Name>**.

Examples

```
<Source_Name>  
  <Short_Name>AQUA</Short_Name>  
  <Long_Name>Earth Observing System, AQUA<Long_Name>  
</Source_Name>
```

```
<Source_Name>  
  <Short_Name>AUVS</Short_Name>  
  <Long_Name>Autonomous Underwater Vehicles</Long_Name>  
</Source_Name>
```

Temporal Coverage

Definition

The **<Temporal_Coverage>** field specifies the start and stop dates during which the data was collected.

<Temporal_Coverage> consists of:

- **<Start_Date>** is the starting date of the data collection.
- **<Stop_Date>** is the ending date of the data collection.

Syntax

```
<Temporal_Coverage>  
  <Start_Date>yyyy-mm-dd</Start_Date>  
  <Stop_Date>yyyy-mm-dd</Stop_Date>  
</Temporal_Coverage>
```

Year, month and day must follow the International Standard ISO 8601 in the form: yyyy-mm-dd, where:

yyyy= the 4 digit year
mm= the 2-digit month (01-12)
dd = the 2-digit day (01-31)

Usage

<Temporal_Coverage> is a highly recommended field and may be repeated.

<Start_Date> and **<Stop_Date>** may not be repeated within

<Temporal_Coverage>.

<Stop_Date> should be omitted if data continue through the present.

<Start_Date> may be omitted if the starting time of data collection is completely unknown, as in a database where data had been collected over a long period of time or had been passed down to various authorities. The inclusion of a

<Stop_Date> without a **<Start_Date>** is not valid in the GCMD DIF Schema.

For map data sets, the publication date should be entered in both **<Start_Date>** and **<Stop_Date>**.

Repeating the **<Temporal_Coverage>** within the metadata record implies that there are gaps in the data coverage and that the temporal coverage is not continuous. This field is not to be used to indicate the temporal resolution of the

data set (see **<Data Resolution>**).

Example

```
<Temporal_Coverage>  
  <Start_Date>2006-12-01</Start_Date>  
  <Stop_Date>2007-02-15</Stop_Date>  
</Temporal_Coverage>
```

Example of temporal coverage that describes data that is continuously being collected (i.e. no <Stop_Date> is necessary).

```
<Temporal_Coverage>  
  <Start_Date>2005-12-01</Start_Date>  
</Temporal_Coverage>
```

Paleo Temporal Coverage

Definition

For paleoclimate or geologic data, **<Paleo_Temporal_Coverage>** is the length of time represented by the data collected. **<Paleo_Temporal_Coverage>** should be used when the data spans time frames earlier than yyyy-mm-dd = 0001-01-01.

<Paleo_Start_Date> is the number of years furthest back in time including units Ga, Ma, ka or ybp.

<Paleo_Stop_Date> is the number of years closest to the present time including units Ga, Ma, ka or ybp.

<Chronostratigraphic_Unit> is a controlled keyword describing Geologic Time of the form:

- **<Eon>** is the primary and largest division of geologic time. Limits correspond with major, global crustal events, changes in sea level and/or climate, or biotic changes.
- **<Era>** is a major division of geological time, tens or hundreds of millions of years long, usually distinguished by significant changes in the plant and animal kingdoms; also used to denote later archaeological periods, such as the prehistoric era. Five geologic eras have been established: Archeozoic [before 2500 million years ago (Ma)], Proterozoic (2500 to 570 Ma), Paleozoic (570 Ma to ca. 250 Ma), Mesozoic (ca. 250 to ca. 70 Ma), and Cenozoic (since ca. 70 Ma). All eras are divided into at least two geologic periods.
- **<Period>** is the secondary division of geologic time, delimited by moderate but usually global crustal events, changes in sea level and/or climate, or biotic changes, sometimes in a relatively localized area. Two or more periods are required to make up a geologic era, and each period comprises two or more geologic epochs.
- **<Epoch>** is the third-order division of geologic time, delimited by partial withdrawal of the sea from land masses and by gentle crustal disturbances in localized areas. Two or more epochs are required to make up a geologic period, and, in turn, two or more periods are needed to constitute a geologic era.
- **<Stage>** is a smaller unit of time. An Epoch usually has 1 or more Stages.

Syntax

```
<Paleo_Temporal_Coverage>  
<Paleo_Start_Date>years [Ga, Ma, ka or ybp]</Paleo_Start_Date>  
<Paleo_Stop_Date>years [Ga, Ma, ka or ybp]</Paleo_Stop_Date>  
<Chronostratigraphic_Unit>Chronostratigraphic Unit keywords  
</Chronostratigraphic_Unit>  
</Paleo_Temporal_Coverage>
```

- The subfields <Paleo_Start_Date>and<Paleo_Stop_Date> may contain up to 80 characters of the UTF-8 character set.

Usage

The **<Paleo_Temporal_Coverage>** field is highly recommended and may be repeated.

<Paleo Start Date> and **<Paleo Stop Date>** are highly recommended subfields and may not be repeated, but when they occur, they must occur together, i.e. both **<Paleo Start Date>** and **<Paleo Stop Date>** must be included.

<Paleo Start Date> and **<Paleo Stop Date>** must include units of time, e.g. Ga, Ma, ka and ybp; where *Ga* = *billion years ago*, *Ma* = *million years ago*, *Ka* = *thousand years ago* and *ybp* = *years before present (present being about 1950)*.

The **<Chronostratigraphic Unit>** subfield is recommended and may be repeated. **<Chronostratigraphic Unit>** must be selected from the **<Chronostratigraphic Unit>** Keywords.

Example

```
<Paleo_Temporal_Coverage>  
<Paleo_Start_Date>5 Ma</Paleo_Start_Date>  
<Paleo_Stop_Date>2 Ma</Paleo_Stop_Date>  
<Chronostratigraphic_Unit>
```

<Eon>PHANEROZOIC</Eon>

<Era>CENOZOIC</Era>

<Period>NEOGENE</Period>

<Epoch>Pliocene</Epoch>

</Chronostratigraphic_Unit>

</Paleo_Temporal_Coverage>

<Paleo_Temporal_Coverage>

<Paleo_Start_Date>2000 ybp</Paleo_Start_Date>

<Paleo_Stop_Date>0 ybp</Paleo_Stop_Date>

</Paleo_Temporal_Coverage>

<Paleo_Temporal_Coverage>

<Paleo_Start_Date>23 Ma</Paleo_Start_Date>

<Paleo_Stop_Date>10 ka</Paleo_Stop_Date>

<Chronostratigraphic_Unit>

<Eon>PHANEROZOIC</Eon>

<Era>CENOZOIC</Era>

<Period>NEOGENE</Period>

<Epoch>MIOCENE</Epoch>

</Chronostratigraphic_Unit>

<Chronostratigraphic_Unit>

<Eon>PHANEROZOIC</Eon>

<Era>CENOZOIC</Era>

<Period>NEOGENE</Period>

<Epoch>PLIOCENE</Epoch>

</Chronostratigraphic_Unit>

<Chronostratigraphic_Unit>

<Eon>PHANEROZOIC</Eon>

<Era>CENOZOIC</Era>

<Period>QUATERNARY</Period>

<Epoch>PLEISTOCENE</Epoch>

</Chronostratigraphic_Unit>

</Paleo_Temporal_Coverage>

Spatial Coverage

Definition

The **<Spatial_Coverage>** field specifies the geographic and vertical (altitude, depth) coverage of the data.

- **<Southernmost_Latitude>** The southernmost geographic latitude covered by the data. From: 0 - 90° North or 0 - 90° South.
- **<Northernmost_Latitude>** The northernmost geographic latitude covered by the data. From: 0 - 90° North or 0 - 90° South.
- **<Westernmost_Longitude>** The westernmost geographic longitude covered by the data. From: 0 - 180° east or 0 - 180° west. The Prime Meridian is 0 degrees, measured positive (+) eastwards of the PM.
- **<Easternmost_Longitude>** The easternmost geographic longitude covered by the data. From: 0 - 180° east or 0 - 180° west. The Prime Meridian is 0 degrees, measured positive (+) eastwards of the PM.
- **<Minimum_Altitude>** The altitude level, which represents the lower limit of data coverage. Units can be specified.
- **<Maximum_Altitude>** The altitude level, which represents the higher limit of data coverage. Units can be specified.
- **<Minimum_Depth>** The depth level, which represents the upper-most depth of data coverage. Units can be specified.
- **<Maximum_Depth>** The depth level, which represents the lowest depth of data coverage. Units can be specified.

Syntax

```
<Spatial_Coverage>  
<Southernmost_Latitude>0 - 90° North or 0 - 90° South</Southernmost_Latitude>  
<Northernmost_Latitude>0 - 90° North or 0 - 90° South</Northernmost_Latitude>  
<Westernmost_Longitude>0 - 180° east or 0 - 180° west</Westernmost_Longitude>  
<Easternmost_Longitude>0 - 180° east or 0 - 180° west</Easternmost_Longitude>  
<Minimum_Altitude>Text up to 80 characters</Minimum_Altitude>  
<Maximum_Altitude>Text up to 80 characters</Maximum_Altitude>  
<Minimum_Depth>Text up to 80 characters</Minimum_Depth>  
<Maximum_Depth>Text up to 80 characters</Maximum_Depth>  
</Spatial_Coverage>
```

- Characters may be selected from the printable ASCII character set.

- A minus sign, "-", placed immediately before the latitude value indicates a latitude south of the Equator (defined as 0 degrees latitude). A minus sign, "-", placed immediately before the longitude value indicates a longitude west of the Prime Meridian (defined as 0 degrees). Latitude north of the Equator (positive) and a longitude east of the Prime Meridian (positive) are assumed if minus signs are not entered before the latitude and longitude values.

N, S, E, or W do not have to be specified, but can be used in place of minus signs. N indicates latitude North of the Equator, S indicates latitude South of the Equator, E indicates longitude East of the Prime Meridian, and W indicates longitude West of the Prime Meridian.

Usage

The **<Spatial_Coverage>** field is highly recommended and may be repeatable.

Each field within **<Spatial_Coverage>** is optional and may not be repeated, with the following exception:

If latitude or longitudes are specified, all four values (**<Northernmost_Latitude>** and **<Southernmost_Latitude>**, **<Westernmost_Longitude>** and **<Easternmost_Longitude>**) must be specified.

Point data can be specified by repeating **<Northernmost_Latitude>** and **<Southernmost_Latitude>** and **<Westernmost_Longitude>** and **<Easternmost_Longitude>** values.

Note:

1. Pay careful attention to coordinates when specifying the bounding box of the data set. Remember that the **<Southernmost_Latitude>** may be in the Northern Hemisphere. Likewise, **<Northernmost_Latitude>** may be in the Southern Hemisphere, **<Westernmost_Longitude>** may be in the Eastern Hemisphere, and **<Easternmost_Longitude>** may be in the Western Hemisphere.
2. If data coverage extends over the North Pole, the **<Northernmost_Latitude>** is 90N (or +90), the **<Southernmost_Latitude>** is the southernmost extent of the data, the **<Easternmost_Longitude>** is 180E (or +180), and the **<Westernmost_Longitude>** is 180W (or -180).
3. If data coverage extends over the South Pole, the

<Southernmost_Latitude> is 90S (or -90), the
<Northernmost_Latitude> is the northernmost extent of the data,
the **<Easternmost_Longitude>** is 180E (or +180), and the
<Westernmost_Longitude> is 180W (or -180).

4. If data coverage extends over the International Date Line (180 ° Longitude), the **<Easternmost_Longitude>** of the data will likely be in the Western Hemisphere, and the **<Westernmost_Longitude>** will likely be in the Eastern Hemisphere.

Some examples of altitude and depth units are 'm' (meters), 'km' (kilometers), 'ft' (feet), 'surface', 'hPA' (hecto Pascals), and 'mb' (millibars).

Example

1. Bounding Box Data Set:

```
<Spatial_Coverage>
  <Southernmost_Latitude>25.0</Southernmost_Latitude>
  <Northernmost_Latitude>50.0</Northernmost_Latitude>
  <Westernmost_Longitude>-100.0</Westernmost_Longitude>
  <Easternmost_Longitude>-50.0</Easternmost_Longitude>
  <Minimum_Altitude>5 miles</Minimum_Altitude>
  <Maximum_Altitude>10 miles</Maximum_Altitude>
  <Minimum_Depth>10 feet</Minimum_Depth>
  <Maximum_Depth>100 feet</Maximum_Depth>
</Spatial_Coverage>
```

2. Global Data Set:

```
<Spatial_Coverage>
  <Southernmost_Latitude>90S</Southernmost_Latitude>
  <Northernmost_Latitude>90N</Northernmost_Latitude>
  <Westernmost_Longitude>180W</Westernmost_Longitude>
  <Easternmost_Longitude>180E</Easternmost_Longitude>
</Spatial_Coverage>
```

3. Point Data Set:

```
<Spatial_Coverage>
  <Southernmost_Latitude>-61</Southernmost_Latitude>
  <Northernmost_Latitude>-61</Northernmost_Latitude>
  <Westernmost_Longitude>-46</Westernmost_Longitude>
  <Easternmost_Longitude>-46</Easternmost_Longitude>
</Spatial_Coverage>
```

Location

Definition

The **<Location>** field specifies 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 under the following **Categories**:

- **Continent**: specifies continents:
 - Africa
 - Antarctica
 - Asia
 - Australia/New Zealand
 - Europe
 - North America
 - South America
- **Ocean**: specifies an ocean body:
 - Arctic Ocean
 - Atlantic Ocean
 - Indian Ocean
 - Pacific Ocean
 - Southern Ocean
- **Geographic Region**: specifies a region on Earth (e.g., Polar)
 - Arctic
 - Eastern Hemisphere
 - Equatorial
 - Eurasia
 - Global
 - Global Land
 - Global Ocean
 - Mid-latitude
 - Northern Hemisphere
 - Oceania
 - Polar
 - Sahel
 - Southern Hemisphere
 - Tropics

- Western Hemisphere
- **Solid Earth:** specifies a region within the Earth:
 - Core
 - Crust
 - Mantle
- **Space:** specifies a region outside of Earth, particularly regions of Earth-Sun interactions:
 - Earth Magnetic Field
 - Solar Region
- **Vertical Location:** specifies a region in the vertical direction, either in the atmosphere or under the oceans:
 - Boundary Layer
 - Ionosphere
 - Land Surface
 - Mesosphere
 - Sea Floor
 - Sea Surface
 - Stratosphere
 - Thermosphere
 - Troposphere

Syntax

```
<Location>
  <Location_Category>Location Category keyword
  </Location_Category>
  <Location_Type>Location type keyword</Location_Type>
  <Location_Subregion1>Location Subregion1 keyword </Location_Subregion1>
  <Location_Subregion2>Location Subregion2 keyword</Location_Subregion2>
  <Location_Subregion3>Location Subregion3 keyword</Location_Subregion3>
  <Detailed_Location>Text</Detailed_Location>
</Location>
```

- **<Detailed_Location>** can contain up to 80 characters.
- Characters may be selected from the UTF-8 character set.

Usage

The **<Location>** field is highly recommended and repeatable.

The **<Location>** subfields are not repeatable within the **<Location>** field.

Location keywords must be selected from the set of controlled Location keywords.

<Detailed_Location> allows authors to include any location keywords that are not available in the controlled keyword list.

Example

```
<Location>
  <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>
```


Data Resolution

Definition

The **<Data_Resolution>** field specifies the resolution of the data, which is the difference between two adjacent geographic, vertical, or temporal values. Controlled keywords representing horizontal, vertical and temporal data resolution ranges can be selected. Selection of data resolution ranges will assist users in refining their search for data within specific resolution ranges.

- **<Latitude_Resolution>** is the minimum difference between two adjacent latitude values. For example, **<Latitude_Resolution>2.5 degrees </Latitude_Resolution>**
- **<Longitude_Resolution>** is the minimum difference between two adjacent longitude values.
- **<Horizontal_Resolution_Range>** is a controlled list of latitude/longitude resolution ranges. The range should be selected based on the **<Latitude_Resolution>** and **<Longitude_Resolution>**.
- **<Vertical_Resolution>** is the minimum difference possible between two adjacent vertical values.
- **<Vertical_Resolution_Range>** is a controlled list of vertical resolution ranges. The range should be selected based on the specified **<Vertical_Resolution>**.
- **<Temporal_Resolution>** is the frequency of data sampled.
<Temporal_Resolution_Range> is a controlled list of temporal resolution ranges. The range should be selected based on the specified **<Temporal_Resolution>**.

Syntax

```
<Data_Resolution>
  <Latitude_Resolution>Text, including units</Latitude_Resolution>
  <Longitude_Resolution>Text, including units</Longitude_Resolution>
  <Horizontal_Resolution_Range>controlled horizontal resolution range keyword
</Horizontal_Resolution_Range>
  <Vertical_Resolution>Text, including units </Vertical_Resolution>
  <Vertical_Resolution_Range>controlled vertical resolution range
keyword</Vertical_Resolution_Range>
  <Temporal_Resolution>Text, including units </Temporal_Resolution>
  <Temporal_Resolution_Range>controlled temporal resolution range keyword
</Temporal_Resolution_Range>
</Data_Resolution>
```

- All fields may contain 1 to 80 characters of the UTF-8 character set.

Usage

The **<Data_Resolution>** field is highly recommended.

The **<Data_Resolution>** field may be repeated.

The **<Data_Resolution>** subfields are optional and may not be repeated within **<Data_Resolution>**.

The author may enter precise horizontal and vertical spatial resolutions, as well as temporal resolutions.

The **<Horizontal_Resolution_Range>**, **<Vertical_Resolution_Range>** and the **<Temporal_Resolution_Range>** keywords should be selected from the set of data resolution keywords.

Units must be specified in the **<Data_Resolution>** text fields. Some suggested units are: m (meters), km (kilometers), degree, hourly, daily, weekly, and monthly.

Example

```
<Data_Resolution>
  <Latitude_Resolution>10 km</Latitude_Resolution>
  <Longitude_Resolution>10 km</Longitude_Resolution>
  <Horizontal_Resolution_Range>10 km - < 50 km or approximately .09 degree - < .5
degree</Horizontal_Resolution_Range>
  <Vertical_Resolution> 100 m </Vertical_Resolution>
  <Vertical_Resolution_Range>30 m -< 100 m </Vertical_Resolution_Range>
  <Temporal_Resolution>Two day repeat observations with a repeat orbit pattern
every 16 days.</Temporal_Resolution>
  <Temporal_Resolution_Range>Daily - < Weekly</Temporal_Resolution_Range>
</Data_Resolution>
```

Project

Definition

The **<Project>** is the name of the scientific program, field campaign, or project from which the data were collected.

This field allows for the specification of keywords that are the names of the projects. These keywords are important for the search and retrieval of information from the GCMD. The field **<Project>** consists of the short name and the long name of the project. There is a 1:1 correspondence between the short name and the long name.

- **<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.

Syntax

```
<Project>  
  <Short_Name>Project short name keyword</Short_Name>  
  <Long_Name>Project long name keyword</Long_Name>  
</Project>
```

Usage

The **<Project>** field is highly recommended and repeatable.

<Short_Name> 1 to 80 characters from the UTF-8 character set, except ">"

<Long_Name> 1 to 220 characters from the UTF-8 character set, except ">"

<Short_Name> and **<Long_Name>** must be selected from the controlled set of Project keywords. However, new project identifiers may only be added if the project does not already appear in the Project keyword List.

The **<Short_Name>** is usually a mnemonic or abbreviated version of the **<Long_Name>**. When submitting a new **<Short_Name>**, always check the Project Keyword Lists to ensure mutual exclusivity among project short names. No two projects may have the same **<Short_Name>**.

Example

```
<Project>  
  <Short_Name>OBIS</Short_Name>  
  <Long_Name>Ocean Biogeographic Information System</Long_Name>  
</Project>
```

```
<Project>  
  <Short_Name>EOS</Short_Name>  
  <Long_Name>Earth Observing System</Long_Name>  
</Project>
```

Quality

Definition

The **<Quality>** field allows the author to provide information about the quality of the data or any quality assurance procedures followed in producing the data described in the metadata. Suggestions for information to include in the **<Quality>** field

- Description should be succinct.
- Include indicators of data quality or quality flags.
- Include recognized or potential problems with quality.
- Established quality control mechanisms should be included.
- Established quantitative quality measurements should be included.

Syntax

<Quality>multiple lines of text**</Quality>**

Multiple lines of text are allowed for the **<Quality>** field and may contain characters from the UTF-8 character set.

Usage

The **<Quality>** field is highly recommended and may not be repeated.

Examples

<Quality>The data set has been in extensive use for over 5 years by many researchers and has resulted in multiple publications.**</Quality>**

<Quality>Orbit calculations using the JGM-3 gravity model have reduced the radial orbit error to 10-15 cm.**</Quality>**

<Quality>This data has been in extensive use in weather forecasting and atmospheric chemistry research for many years, and has thus been subject to independent review by many scientific experts.**</Quality>**

Access Constraints

Definition

The **<Access_Constraints>** field allows the author to provide information about any constraints for accessing the data set. This includes any special restrictions, legal prerequisites, limitations and/or warnings on obtaining the data set. Some words that may be used in this field include: Public, In-house, Limited, Additional detailed instructions on how to access the data can be entered in this field.

Syntax

`<Access_Constraints>text</Access_Constraints>`

- Multiple lines of text are permitted. Characters may be selected from the UTF-8 character set.

Usage

The **<Access_Constraints>** field is highly recommended and may not be repeated. Text should be single spaced with blank lines separating paragraphs

Example

`<Access_Constraints>Registration is required to access the data.</Access_Constraints>`

Use Constraints

Definition

The **<Use_Constraints>** field allows the author to describe how the data may or may not be used after access is granted to assure the protection of privacy or intellectual property. This includes any special restrictions, legal prerequisites, terms and conditions, and/or limitations on using the data set. Data providers may request acknowledgement of the data from users and claim no responsibility for quality and completeness of data.

Syntax

`<Use_Constraints>text</Use_Constraints>`

- Multiple lines of text are permitted. Characters may be selected from the UTF-8 character set.

Usage

The **<Use_Constraints>** field is highly recommended and may not be repeated. Text should be single spaced with blank lines separating paragraphs.

Example

`<Use_Constraints>`

Acknowledge the use of specific records from contributing databases in the form appearing in the 'Citation' field thereof (if any); and acknowledge the use of the OBIS facility.

For information purposes, email to [obissupport at marine.rutgers.edu](mailto:obissupport@marine.rutgers.edu) the full citation of any publication made (printed or electronic) that cites OBIS or any constituent part.

Recognize the limitations of data in OBIS.

`</Use_Constraints>`

`<Use_Constraints>`

Data may not be used for commercial applications.

`</Use_Constraints>`

Distribution

Definition

The **<Distribution>** field describes media options, size, data format, and fees involved in distributing the data set.

- **<Distribution_Media>** The media options for the user receiving the data.
- **<Distribution_Size>** An approximate size (in KB, MB or GB) for the entire data set. Specify if data are compressed and the method of compression.
- **<Distribution_Format>** The data format used to distribute the data.
- **<Fees>** Cost of **<Distribution_Media>** or distribution costs if any. Specify if there are no costs.

Syntax

```
<Distribution>
  <Distribution_Media>Text up to 80 characters</Distribution_Media>
  <Distribution_Size>Text up to 80 characters</Distribution_Size>
  <Distribution_Format>Text up to 80 characters </Distribution_Format>
  <Fees> Text up to 80 characters</Fees>
</Distribution>
```

- Characters may be selected from the UTF-8 character set.

Usage

The **<Distribution>** field is highly recommended and may be repeated.

<Distribution> subfields may only appear once within **<Distribution>**.

Keywords for **<Distribution_Media>** may be selected from the Suggested Distribution Media Keywords List.

Keywords for **<Distribution_Format>** may be selected from the Suggested Distribution Format Keywords List.

Example

```
<Distribution>
  <Distribution_Media>CD-ROM</Distribution_Media>
```


<Distribution_Size>25 MB</Distribution_Size>
<Distribution_Format>ESRI shapefile</Distribution_Format>
<Fees>\$20</Fees>
</Distribution>

Data Set Language

Definition

<Data_Set_Language> describes the language used in the preparation, storage, and description of the data. It is the language of the information object, not the language used to describe or interact with the metadata record. **<Data_Set_Language>** does not refer to the language of the metadata.

Syntax

`<Data_Set_Language>text</Data_Set_Language>`

- Consists of 1 - 80 characters from the UTF-8 character set.

Usage

The **<Data_Set_Language>** field is highly recommended and can be repeated if data is available in more than one language.

<Data_Set_Language> may be selected from the Suggested Language Keywords List, which is adapted from the ISO 639 language codes:

http://www.loc.gov/standards/iso639-2/php/code_list.php

Examples

`<Data_Set_Language>English</Data_Set_Language>`

`<Data_Set_Language>French</Data_Set_Language>`

Data Set Progress

Definition

The **<Data_Set_Progress>** describes the production status of the data set regarding its completeness. There are three choices:

- **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.

Syntax

<Data_Set_Progress>Data Set Progress keyword (see above)**</Data_Set_Progress>**

- Up to 31 characters from the UTF-8 character set are allowed.

Usage

The **<Data_Set_Progress>** field is highly recommended and may not be repeated.

Examples

<Data_Set_Progress>Complete**</Data_Set_Progress>**

<Data_Set_Progress>In Work**</Data_Set_Progress>**

<Data_Set_Progress>Planned**</Data_Set_Progress>**

Related URL

Definition

The **<Related_URL>** field specifies links to Internet sites that contain information related to the data, as well as related Internet sites such as project home pages, related data archives/servers, metadata extensions, online software packages, web mapping services, and calibration/validation data. The **<Related_URL>** field consists of

- **<URL_Content_Type>** Describes the type of URL being referenced. The **<URL_Content_Type>** is selected from a list of controlled URL content type keywords and consists of:
 - **<Type>** Describes the type resources being referenced by the URL.
 - **<Subtype>** Describes the subtype of the resource being referenced by the URL.
- **<URL>** is the URL to the resource associated with the data set.
- **<Description>** provides information about the resource defined by the **<URL>**.

Syntax

```
Related_URL>
  <URL_Content_Type>
    <Type>Related URL Type keyword</Type>
    <Subtype>Related URL Subtype keyword</Subtype>
  </URL_Content_Type>
  <URL>URL up to 600 characters</URL>
  <Description>Text</Description>
</Related_URL>
```

- **<Description>** may contain multiple lines of text. Characters may be selected from the UTF-8 character set.

Usage

The **<Related_URL>** field is highly recommended and may be repeated.

The **<URL_Content_Type>** and **<URL>** fields are required.

The **<URL_Content_Type>** subfields **<Type>** and **<Subtype>** must be selected from the set of URL Content Type keywords. **<Subtype>** is not required.

The **<URL>** subfield may be repeated within the **<Related_URL>**.

The **<Description>** subfield is recommended and unrepeatable.

Examples

```
<Related_URL>
  <URL_Content_Type>
    <Type>VIEW PROJECT HOME PAGE </Type>
  </URL_Content_Type>
  <URL>http://www.esipfed.org/</URL>
  <Description>Homepage of the Earth Science Information Partners (ESIP)
  </Description>
</Related_URL>
```

Example of a **<Related_URL>** with access to data:

```
<Related_URL>
  <URL_Content_Type>
    <Type> GET DATA </Type>
    <Subtype>GIOVANNI</Subtype>
  </URL_Content_Type>
  <URL> ftp://g0dup05u.ecs.nasa.gov/data/Giovanni/MODIS/GOCART/</URL>
  <Description>Access to data files </Description>
</Related_URL>
```

Example of a **<Related_URL>** with repeating **<URL>** subfields.

```
<Related_URL>
  <URL_Content_Type>
    <Type>GET DATA </Type>
  </URL_Content_Type>
  <URL>ftp://ftp.ncdc.noaa.gov/pub/data/paleo/paleocean/by_contributor/medina-
elizalde2005/medina-elizalde2005.txt </URL>
  <URL>ftp://ftp.ncdc.noaa.gov/pub/data/paleo/paleocean/by_contributor/medina-
elizalde2005/</URL>
  <Description>Access to data and documentation files </Description>
</Related_URL>
```

DIF Revision History

Definition

The **<DIF_Revision_History>** allows the author to provide a list of changes made to the DIF over time. This provides a mechanism for tracking revisions to DIF content.

Syntax

<DIF_Revision_History>yyyy-mm-dd, text of revision history**</DIF_Revision_History>**

Year, month and day must follow the International Standard ISO 8601 in the form: yyyy-mm-dd, where:

yyyy= the 4 digit year
mm= the 2-digit month (01 -12)
dd = the 2-digit day (01-31)

Usage

The **<DIF_Revision_History>** field is recommended and may not be repeated. Multiple lines of text are allowed. Each line may contain up to 600 characters of the UTF-8 character set and should be preceded by the date the change was made.

Example

<DIF_Revision_History>2004-09-05, Jane Doe, DIF author, changed PI telephone number
</DIF_Revision_History>

<DIF_Revision_History>2005-12-17, global update, changed Data_Center_URL 2006-05-22, populated new fields for version 6
</DIF_Revision_History>

Keyword (Ancillary Keyword)

Definition

The **<Keyword>** field allows authors to provide any words or phrases needed to further describe the data set.

Syntax

<Keyword>text up to 160 characters**</Keyword>**

- The **<Keyword>** field may contain up to 160 characters from the UTF-8 character set.

Usage

The **<Keyword>** field is recommended and may be repeated.

Ancillary keywords that are not found in the controlled keyword lists may be provided.

Example

<Keyword>Aerosol Optical Depth**</Keyword>**
<Keyword>Ceiling Height**</Keyword>**
<Keyword>Diffuse Radiation**</Keyword>**
<Keyword>Extraterrestrial Radiation**</Keyword>**
<Keyword>Global Radiation**</Keyword>**
<Keyword>Kentucky**</Keyword>**
<Keyword>Sargasso Sea**</Keyword>**

Originating Center

Definition

The **<Originating_Center>** is the data center or data producer who originally generated the dataset.

Syntax

<Originating_Center>Text</Originating_Center>

- The **<Originating_Center>** field may contain 1 to 240 characters from the UTF-8 character set.
- An acronym to identify the originating data center or data producer is sufficient.

Usage

The **<Originating_Center>** is recommended and may not be repeated.

Example

<Originating_Center>GES-DISC DAAC</Originating_Center>

Multimedia Sample

Definition

The **<Multimedia_Sample>** field allows the author to provide information that will enable the display of a sample image, movie or sound clip within the DIF.

- **<File>** describes the filename where the multimedia sample can be found. If the multimedia sample file resides on the GCMD server, the filename should be specified in this field. If the file is to be transferred to the GCMD, arrangements should be made with GCMD staff.(Note: this is not typical).
- **<URL>** is the URL to be accessed .
- **<Format>** is the format of the multimedia sample file (which may differ from the format of the data files), i.e. GIF, TIFF, JPEG.
- **<Caption>** is a one-line description of the multimedia sample used as a caption when the sample is displayed. The caption is especially useful for images such as graphs and photos.
- **<Description>** A more detailed description of the multimedia sample.

Syntax

```
<Multimedia_Sample>
  <File>text, up to 80 characters</File>
  <URL>URL up to 600 characters</URL>
  <Format>text up to 80 characters</Format>
  <Caption>text up to 80 characters</Caption>
  <Description>multiple lines of text</Description>
</Multimedia_Sample>
```

- Each field may contain one line of 1 to 80 characters, except **<URL>** fields, which may contain one line of 1 to 600 characters. The **<Description>** field may contain multiple lines of 1 to 80 characters per line. Characters may be selected from the UTF-8 character set.

Usage

The **<Multimedia_Sample>** field is recommended and may not be repeated. Individual subfields may not be repeated.

If the **<Multimedia_Sample>** is populated in the DIF, a URL is required.

Keywords for the Format subfield may be selected from the Suggested Multimedia Sample Format Keywords List.

Example

```
<Multimedia_Sample>  
<URL>http://www2.nature.nps.gov/geology/geologic_wonders/images/Tmud_volcano  
.jpg</URL>  
  <Format>GIF Image</Format>  
  <Description> Sample browse graphic of the data set. </Description>  
</Multimedia_Sample>  
  
<Multimedia_Sample>  
  <URL>http://www.longpaddock.qld.gov.au/RainfallAndPastureGrowth/Aus/Rainfall_  
Total.gif</URL>  
  <Format>GIF</Format>  
</Multimedia_Sample>
```

Reference (Publications/References)

Definition

The **<Reference>** field describes key bibliographic citations pertaining to the data set. Bibliographic citations may be provided in styles used by professional scientific journals such as APA or MLA.

Syntax

<Reference>text</Reference>

Usage

The **<Reference>** field is recommended and may not be repeated.
Multiple lines of text are allowed. Characters may be selected from the UTF-8 character set.

Example

<Reference>

Crum, T.D., and R.L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. Bull. Amer. Meteor. Soc., 74, 1669-1685.

Crum, T.D., R.L. Alberty, and D.W. Burgess, 1993: Recording, Archiving, and Using WSR-88D Data. Bull. Amer. Meteor. Soc., 74, 645-653.

Klazura, G.E., and D.A. Imy, 1993: A Description of the Initial Set of Analysis Products Available from the NEXRAD WSR-88D System. Bull. Amer. Meteor. Soc., 74, 1293-1311.

Cooper, A.K., 1995, Strides made in understanding Antarctic stratigraphy and Cenozoic glacial history, Eos Transactions AGU, March 7, 1995, p.97.

URL; <http://earth.agu.org/kosmos/cooper.html>

</Reference>

Parent DIF

Definition

The **<Parent_DIF>** field allows the capability to relate generalized aggregated metadata records (parents) to metadata records with highly specific information (children).

Population of the **<Parent_DIF>** field should be reserved for instances where many metadata records are basically subsets that can be better represented by one parent metadata record, which describes the entire collection. Typically, the parent metadata record will have many children metadata records, which refer to the parent through the **<Parent_DIF>** field. In some instances, a child may point to more than one parent. The **<Parent_DIF>** is populated with an **<Entry_ID>**.

The **<Parent_DIF>** is only populated if there are children metadata records that reference a parent metadata record. Upon display of results, both parent and children metadata are available with explicit linkages between parent and child records through the **<Parent_DIF>** field.

Syntax

<Parent_DIF>Entry_ID</Parent_DIF>

- The syntax of the **<Parent_DIF>** is identical to the **<Entry_ID>**. The **<Parent_DIF>** consists of 1 to 80 alphanumeric characters of the UTF-8 set, including underbar (_), hyphen (-) and period (.)
- The following characters are not allowed in the **<Parent_DIF>**:
 - Forward Slash (/)
 - Backward Slash (\)
 - Spaces ()
 - Colon (:)

Usage

The **<Parent_DIF>** field is recommended and may be repeated.

The **<Parent_DIF>** field must be populated with the existing **<Entry_ID>** of the parent metadata record.

The field may be repeated in cases where a child metadata record has two different immediate parents, i.e. the child is a member of two distinct aggregations.

Examples

<Parent_DIF>NOAA_HISTORICAL_CLIMATE_DATA</Parent_DIF>

<Parent_DIF>USGS_COASTALDATA_PARENTDIF</Parent_DIF>

IDN Node

Definition

The Internal Directory Name (IDN) Node (**<IDN_Node>**) field is used internally to identify association, responsibility and/or ownership of the dataset, service or supplemental information. Note: The **<IDN_Node>** field is usually not displayed to the user. The author may populate **<IDN_Node>** from a set of controlled keywords.

Syntax

<IDN_Node><Short_Name>IDN Node Keyword</Short_Name></IDN_Node>

Usage

The **<IDN_Node>** field is recommended and may be repeated.

The **<IDN_Node>** must be selected from the set of controlled keywords.

Example

<IDN_Node><Short_Name>USA/NASA</Short_Name></IDN_Node>

DIF Creation Date

Definition

The **<DIF_Creation_Date>** specifies the date the metadata record was created. The **<DIF_Creation_Date>** is automatically populated by the authoring tools with the date the metadata entry was created, however, the author may change the date.

Syntax

<DIF_Creation_Date>yyyy-mm-dd**</DIF_Creation_Date>**

Year, month and day must follow the International Standard ISO 8601 in the form: yyyy-mm-dd, where:

yyyy= the 4 digit year
mm= the 2-digit month (01 -12)
dd = the 2-digit day (01-31)

Usage

The **<DIF_Creation_Date>** is recommended and may not be repeated.

Example

<DIF_Creation_Date>2006-06-17**</DIF_Creation_Date>**

Last DIF Revision Date

Definition

The **<Last_DIF_Revision_Date>** specifies the date the metadata record was created. The **<Last_DIF_Revision_Date>** is automatically populated by the authoring tools with the date the metadata entry was created or modified, however, the author may change the date.

Syntax

<Last_DIF_Revision_Date>yyyy-mm-dd**</Last_DIF_Revision_Date >**

Year, month and day must follow the International Standard ISO 8601 in the form: yyyy-mm-dd, where:

yyyy= the 4 digit year
mm= the 2-digit month (01 -12)
dd = the 2-digit day (01-31)

Usage

The **<Last_DIF_Revision_Date>** is recommended and may not be repeated.

Example

<Last_DIF_Revision_Date>2006-11-01**</Last_DIF_Revision_Date>**

Future DIF Review Date

Definition

The **<Future_DIF_Revision_Date>** allows for the specification of a future date at which the DIF should be reviewed for accuracy of scientific or technical content. Future metadata changes may be due to:

- Anticipated changes in the personnel fields.
- Planned changes to the data content (e.g., change in processing algorithm).
- Expected inclusion of the data as part of a campaign.
- Planned review of metadata and/or data by the data producer.

Syntax

<Future_DIF_Revision_Date>yyyy-mm-dd**</Future_DIF_Revision_Date>**

Year, month and day must follow the International Standard ISO 8601 in the form: yyyy-mm-dd, where:

yyyy= the 4 digit year
mm= the 2-digit month (01 -12)
dd = the 2-digit day (01-31)

Usage

The **<Future_DIF_Revision_Date>** is recommended and may be repeated.

Examples

<Future_DIF_Revision_Date>2007-12-17**</Future_DIF_Revision_Date>**

<Future_DIF_Revision_Date>2008-12-17**</Future_DIF_Revision_Date>**

Privacy Status

Definition

The **<Private>** field allows the author to restrict the data set description from being publicly available. There are only two options:

True. Make the description unavailable to the public.

False. Make the description publicly available. This is the default. If the author wants the description to be available to the public, then no action is required.

Syntax

`<Private>True or False</Private>`

Usage

The **<Private>** field is recommended and may not be repeated.

The **<Private>** flag is always set to **False**. If the author wants to restrict the description from being accessible to the public, then the **<Private>** must be set to **True**.

Example

`<Private>True</Private>`