



Global Change Master Directory (GCMD)/International Directory Network (IDN) DIFs for CEOS WGISS Integrated Catalog (CWIC) Data Provider Products

1. Introduction

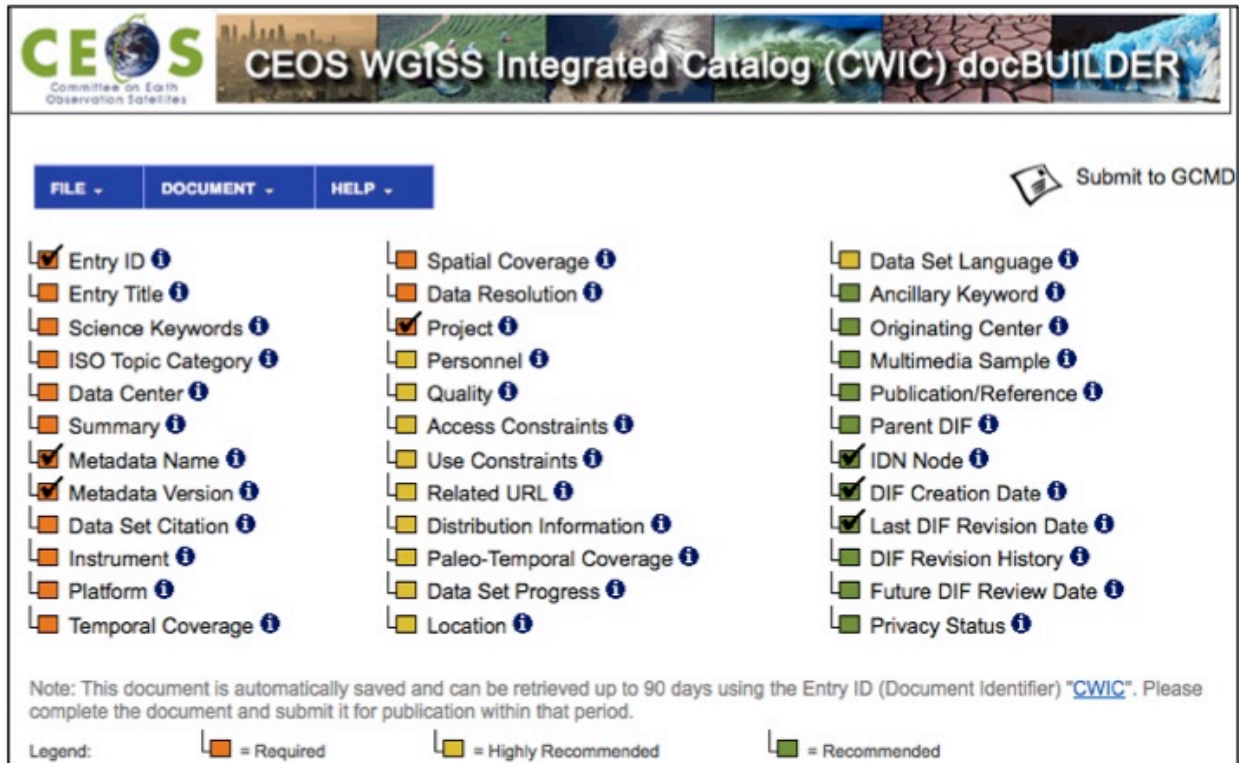
The Directory Interchange Format (DIF) is a metadata format used to create directory entries that describe scientific data sets. Because of the widespread dissemination and use of DIFs, the term, “DIF” also has come to mean discovery level metadata. A DIF holds a collection of fields, which detail specific information about the data set. Seventeen are free-text fields, while nine require the use of controlled keywords. These keywords are maintained within the GCMD database, where validation for mandatory fields and keywords takes place. These controlled keywords are important because they provide normalized searches for users. The DIF offers the essential core data discovery fields, which are required in the ISO 19115 metadata standard, along with these valuable controlled keyword search options. Using the DIF metadata creation tool (docBuilder), discovery level metadata can be entered using the DIF format. Each DIF entry can be extracted in other compatible formats such as the ISO 19115. For CWIC data partners, a DIF for each dataset that has an inventory is required. CWIC portals can search the IDN for high level dataset information and get results returned in the ISO 19115 format.

IDN Home Page: <http://idn.ceos.org/>

Add Data Set Descriptions to the GCMD/IDN: <http://gcmd.nasa.gov/User/authoring.html>

Direct link to the CWIC docBuilder: <http://gcmd.nasa.gov/DocumentBuilder/Home.do?Portal=cwic>

The above links offer information about accessing a user account for using the GCMD/IDN DIF docBuilder tool, online docBuilder guide, and DIF user guides.



(Image of the GCMD/IDN CWIC docBuilder tool)

For information about CWIC:

http://www.ceos.org/index.php?option=com_content&view=category&layout=blog&id=154&Itemid=225

2. GCMD/IDN Fields required by CWIC

The CWIC provides mediated data access to granules in datasets that have been designated as CWIC datasets in the IDN via a specified Project Short Name value of “CWIC”. In order for the CWIC inventories to unambiguously map the IDN dataset to provider system datasets, some of the IDN fields that are normally not required must be provided in a consistent format in order to support the CWIC interaction with the data provider inventory system. The table provided here can be used to guide the author through the DIF fields using any of the available DIF authoring tools. In the table below, the terms in orange indicate required fields. Terms in yellow indicate highly recommended fields. Recommended fields are displayed in green. This “required”, “highly recommended”, and “recommended” color coding in the first column of the table refer to generic DIFs. The third column in the table below identifies which fields are “required” for CWIC accessible datasets. CWIC data providers are strongly encouraged to provide a rich set of metadata in the DIFs to facilitate and support the CWIC interaction with the provider system.

(All Field Names are hyperlinked to the DIF User Guide pages.)

<u>Field:</u>	<u>Definition:</u>	<u>Required by CWIC</u>
<u>Entry ID:</u> *	The "Entry_ID" is the unique document identifier of the metadata record.	Y
<u>Entry Title:</u>	The "Entry_Title" is the title of the data set described by the metadata.	Y
<u>Parameters (Science Keywords):</u>	The "Parameters" field allows for the specification of Earth science keywords that are representative of the data set being described.	Y
<u>ISO Topic Category:</u>	The "ISO_Topic_Category" field is used to identify the keywords in the ISO 19115 - Geographic Information Metadata Topic Category Code List.	(auto-populated)
<u>Data Center:</u>	The "Data Center" is the data center, organization, or institution responsible for distributing the data.	Y
<u>Summary:</u>	The "Summary" field provides a brief description of the data set along with the purpose of the data.	Y
<u>Metadata Name:</u>	The "Metadata_Name" is used to identify the current DIF standard name. This field is auto-populated in docBUILDER.	(auto-populated)
<u>Metadata Version:</u>	The "Metadata_Version" is used to identify the current DIF metadata standard. This field is auto-populated in docBUILDER.	(auto-populated)
<u>Data Set Citation:</u>	The "Data_Set_Citation" field allows the author to properly cite the data set producer.	Y
<u>Data Set Originator/Creator</u>	The name of the organization(s) or individual(s) with primary intellectual responsibility for the data set's development.	Y
<u>Data Set Title</u> **	The title of the data set; this may be the same as Entry Title.	Y
<u>Data Set Release Date</u>	The date when the data set was made available for release.	Y
<u>Data Set Version</u> ***	The version of the data set.	Y
<u>Personnel:</u>	"Personnel" defines the point of contact for more information about the data set or the metadata.	
<u>Related URL:</u>	The "Related_URL" field specifies links to Internet sites that contain information related to the data.	

<u>Instrument (Sensor Name):</u>	The "Instrument" or "Sensor_Name" is the name of the instrument used to acquire the data.	Y
<u>Platform (Source Name):</u>	The "Platform" or "Source_Name" is the name of the platform used to acquire the data.	Y
<u>Temporal Coverage:</u>	The "Temporal_Coverage" field specifies the start and stop dates during which the data were collected.	Y
<u>Paleo-Temporal Coverage:</u>	For paleoclimate or geologic data, "Paleo_Temporal_Coverage" is the length of time represented by the data collected.	
<u>Spatial Coverage:</u>	The "Spatial_Coverage" field specifies the geographic and vertical (altitude, depth) coverage of the data.	Y
<u>Location:</u>	The "Location" field specifies the name of a place on Earth which the data are collected.	
<u>Data Resolution:</u>	The "Data_Resolution" field specifies the resolution of the data, which is the difference between two adjacent geographic, vertical, or temporal values.	Y
<u>Project:****</u>	The "Project" is the name of the scientific program, field campaign, or project from which the data were collected.	Y
<u>Quality:</u>	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.	
<u>Access Constraints:</u>	The "Access_Constraints" field allows the author to provide information about any constraints for accessing the data set.	
<u>Use Constraints:</u>	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.	
<u>Distribution:</u>	The "Distribution" field describes media options, size, data format, and fees involved in distributing the data set.	
<u>Data Set Language:</u>	"Data_Set_Language" describes the language used in the preparation, storage, and description of the data.	
<u>Data Set Progress:</u>	The "Data_Set_Progress" describes the production status of the data set regarding its completeness.	

<u>DIF Revision History:</u>	The “DIF_Revision_History” allows the author to provide a list of changes made to the DIF over time.	
<u>(Ancillary) Keyword:</u>	The “Keyword” field allows authors to provide any words or phrases needed to further describe the data set.	
<u>Originating Center:</u>	The “Originating_Center” is the data center or data producer who originally generated the dataset.	
<u>Multimedia Sample:</u>	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.	
<u>References/Publications:</u>	The “Reference” field describes key bibliographic citations pertaining to the data set.	
<u>Parent DIF:</u>	The “Parent_DIF” field allows the capability to relate generalized aggregated metadata records (parents) to metadata records with highly specific information (children).	
<u>IDN Node:</u>	The Internal Directory Name (IDN) Node field is used internally to identify association, responsibility and/or ownership of the dataset, service or supplemental information.	
<u>DIF Creation Date:</u>	The “DIF_Creation_Date” specifies the date the metadata record was created.	
<u>Last DIF Revision Date:</u>	The “Last_DIF_Revision_Date” specifies the date the metadata record was created or last modified.	
<u>Future DIF Revision Date:</u>	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.	
<u>Privacy Status:</u>	The “Private” field allows the author to restrict the data set description from being publicly available.	

- * Entry_ID should be the same as the Data Set ID.
- ** Data Set Title must be identical to the value that is registered at the data granule archive center.
- *** Data Set Version should be the same integer value as in the data granule metadata files.
- **** Project keyword should be set to ‘CWIC’.

3. DIF Examples:

IRMSS - Infrared Multispectral Scanner (CBERS 2) Imagery

http://gcmd.gsfc.nasa.gov/getdif.htm?INPE_CBERS2_IRM

Landsat Thematic Mapper Imagery

http://gcmd.gsfc.nasa.gov/getdif.htm?LANDSAT_TM

MODIS/Terra Calibrated Radiances 5-Min L1B Swath 1km (MOD021KM, Collection 004 and 005)

<http://gcmd.gsfc.nasa.gov/getdif.htm?MOD021KM>

4. GCMD Contacts

If a user requires an account or has any questions, please contact the GCMD User Support Office at (gsfc-gcmduso@mail.nasa.gov) or go to the User Support link at <http://gcmd.nasa.gov/MailComments/MailComments.jsf?rcpt=gcmduso>.

This document should be cited as:

Directory Interchange Format (DIF) Writer's Guide, 2012. Global Change Master Directory. National Aeronautics and Space Administration. [<http://gcmd.nasa.gov/User/difguide/>]