

DAACs and Distributors

Distributed Active Archive Centers (DAACs) and Distributors

Describing how to access data is an important role of metadata as data access follows data discovery for many users. Once an interesting data set is discovered, users want to get the data so they can use it. This access information is usually provided as descriptions of a DAAC or a Distributor, contact information and a set of links to the data.

NASA GCMD Directory Interchange Format

Organizations

Organizations (and contacts) that are responsible for distributing data are "[Data Centers](#)" in the Directory Interchange Format. They have the following properties:

```
<dif:Data_Center uuid="UUID">
  <dif:Data_Center_Name/>
  <dif:Data_Center_URL/>
  <dif:Data_Set_ID/>
  <dif:Personnel/>
</dif:Data_Center>
```

The Data_Center field is required and can be repeated.

Online Resources

URL's where data can be obtained are in the [Related_URL](#) section of the metadata record. They have the following properties:

```
<dif:Related_URL uuid="UUID">
  <dif:URL_Content_Type/>
  <dif:URL/>
  <dif:Description/>
</dif:Related_URL>
```

Data access URLs can be recognized by the URL_Content_Type = "GET DATA". The Related_URL field is highly recommended and may be repeated.

Download Details

Details about the downloads are included in the [Distribution](#) section of the metadata record. They include:

```
<dif:Distribution>
  <dif:Distribution_Media/>
  <dif:Distribution_Size/>
  <dif:Distribution_Format/>
  <dif:Fees/>
</dif:Distribution>
```

The Distribution field is highly recommended and may be repeated.

ECHO

The ECHO metadata model includes several data access elements.

Organizations

```
<echo:ArchiveCenter/>
<echo:Contact>
  <echo:Role = "distributor"/>
  <echo:HoursOfService/>
  <echo:Instructions/>
  <echo:OrganizationName/>
  <echo:OrganizationAddresses/>
  <echo:OrganizationPhones/>
  <echo:OrganizationEmails/>
  <echo:ContactPersons/>
</echo:Contact>
```

OnlineResources

```

<echo:OnlineAccessURL>
  <echo:URL/>
  <echo:URLDescription/>
  <echo:MimeType/>
</echo:OnlineAccessURL>

```

ISO 19115

Data distribution information is described in "[gmd:MD_Distribution](#)" sections in ISO 19115. They have the following properties:

```

<gmd:distributionInfo>
  <gmd:MD_Distribution>
    <gmd:distributionFormat/>
    <gmd:distributor/>
    <gmd:transferOptions/>
  </gmd:MD_Distribution>
</gmd:distributionInfo>

```

Connections

The information content of these models for distribution information overlaps significantly so it is possible to represent most important content in all three dialects and to do some translations without losing information. There are a few differences that might be important in specific cases:

DIF to ISO

1. The DIF model separates information about organizations (Data Centers) from information about people that work in these organizations (Personnel). The ISO model combines organizations, positions, and people into a single object ([gmd:CI_ResponsibleParty](#) or [gmd:CI_Party](#) in 19115-1). The information in the [dif:Data_Center](#) object is combined with information from the [dif:Data_Center/dif:Personnel](#) object into a single [gmd:distributorContact](#) object in the ISO model.
2. The DIF model assigns roles to people and provides a list of standard role names (Investigator, Technical Contact, or DIF Author). The ISO model has a longer list of [role names](#) that includes distributor. In the translation from DIF to ISO, the [dif:Personnel/dif:Role](#) translates to [gmd:positionName](#) in order to preserve the DIF information as well as the standard ISO code.
3. The [Data_Center_Name](#) includes a ShortName and LongName that must be selected from the GCMD [Data Center](#) Keyword list. A decision must be made how these should be combined in the ISO [gmd:organizationName](#). Currently the combination is [gmd:organizationName](#) = [dif:ShortName](#) > [dif:LongName](#).
4. The [dif:Data_Set_ID](#) included in the [dif:Data_Center](#) object identifies the dataset and is controlled by the Data Center. The inclusion of this attribute in the [dif:Data_Center](#) object makes it difficult to reuse Data Center information across multiple records and may make it more difficult to translate this information into different dialects. The translation to ISO separates the identifier from the distribution information. It becomes an [gmd:MD_Identifier](#) for the data set with an authority of the Data Center.
5. The [dif:Distribution_Format](#) field holds a format name from the suggested [Format Keywords](#) list. The ISO [gmd:MD_Format](#) object includes a name that can match the DIF keyword along with a Version, a reference to the specification (19115-1), and other information.
6. The [dif:Distribution_Media](#) holds a media name from the suggested [Media Keywords](#). The ISO [gmd:MD_Medium](#) object includes a [gmd:mediumName](#) attribute that can hold the [dif:distributionMedia](#) information along with a [MD_MediumFormat codeList](#) that provides a shared vocabulary for medium formats, and other information.
7. Many-to-Many relationships: All three of the distribution elements ([dif:Data_Center](#), [dif:Distribution](#), and [dif:Related_URL](#)) are repeatable in a DIF record. If there are more than one of any of these elements the relationships between them may not be clear. For example, if there are two [dif:Distribution/dif:Distribution_Format](#) objects and two [dif:Data_Center](#) objects, it is not clear how one could tell which format is distributed by which [dif:Data_Center](#) or, if there are more than one GET DATA URLs, which format is available from which URL. These relationships are clear in the ISO model because related information is grouped in a [distributionInfo](#) or [distributor](#) object.

ECHO to ISO

1. The [echo:ArchiveCenter](#) is typically the same as the [echo:OrganizationName](#) in the [echo:Contact](#) with the [echo:Role](#) = "distributor".