

Data Discovery



- 1.Data discovery systems
- 2.TDS and metadata
- 3.NcISO services

1.1 GI-Cat, Geoportal, GeoNetwork

- TDS does not support searching features.
- We look for data browsing through catalogs.
- Data discovery systems harvest metadata from TDS catalogs and expose it through interfaces that TDS itself does not support (Brokers).
- Data discovery systems enable searches on TDS servers!
- Harvest data from other sources (OGC services, ISO metadata standards), federated systems...

1.2 Issues harvesting metadata

- TDS instances don't provide too much metadata.
- Crawl at dataset level. Don't harvest metadata at collection level.
- Manage large numbers of datasets.
- Provide access to datasets rather than to actual data.
- Data display that is useful, meaningful and satisfactory for multi-disciplinary users.

1.3 GI-Cat & Geoportal demo

Running instances:

* EURO GEOSS (GI-Cat):

<http://www.eurogeoss-broker.eu/>

* NGDC geoportal:

<http://www.ngdc.noaa.gov/geoportal/catalog/main/home.page>

* NODC geoportal:

<http://data.nodc.noaa.gov/geoportal/catalog/search/search.page>

2.1 Metadata matters



- **METADATA ENABLES DATA DISCOVERY**
- **BETTER INTEROPERABILITY**
- **IMPROVES DATA EXCHANGE**

2.2 We are data providers...

Catalog http://localhost:8081/thredds/metadata_catalog.html

Dataset: photo

- *Data type:* IMAGE
- *ID:* photoSharing

Access:

1. **HTTPServer:** </thredds/fileServer/testStandardTds/images/22.jpg>

Viewers:

- [NetCDF-Java ToolsUI \(webstart\)](#)

Catalog http://localhost:8081/thredds/metadata_catalog.html

Dataset: photo

- *Data format:* image/jpeg
- *Data size:* 29.0 Kbytes
- *Data type:* IMAGE
- *ID:* photoSharing

Documentation:

- **summary:** This is my staff picture taken few weeks after I started to work at UNIDATA
- **rights:** This image is distributed under creative commons license

Access:

1. **HTTPServer:** </thredds/fileServer/testStandardTds/images/22.jpg>

Contributors:

- Marcos Hermida (**Software Engineer**)

Keywords:

- staff pictures

Projects:

- UNIDATA Training Workshop 2012

Creators:

- **Marcos Hermida**
 - *email:* mhermida@unidata.ucar.edu
 - *http:* <http://www.unidata.ucar.edu>

Properties:

- category = "staff pictures"

Viewers:

- [NetCDF-Java ToolsUI \(webstart\)](#)

2.3 TDS metadata

- TDS provides several kinds of metadata: access, spatial and time coverage, contact, title, summary, keywords, variables....
- Feature Collections get metadata from files (CF-conventions, GRIB...)
- Attribute Conventions Dataset Discovery (ACDD)
[http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery_\(ACDD\)](http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery_(ACDD))

2.4 Adding metadata to catalogs

- Metadata tags:
 - Documentation, metadata, property...
 - Reference:

<http://www.unidata.ucar.edu/projects/THREDDS/tech/catalog/v1.0.2/InvCatalogSpec.html#metadataElement>

- Inheritance: inherited="true"
- ACDD recommendations:

[http://wiki.esipfed.org//index.php/Attribute_Convention_for_Data_Discovery_\(ACDD\)](http://wiki.esipfed.org//index.php/Attribute_Convention_for_Data_Discovery_(ACDD))

3.1 What is NcISO?

- Developed by NGDC (National Geophysical Data Center)
- NcISO provides a set of services for managing our metadata:
 - NCML: information of our datasets in ncml
 - UDDC(*): assesses our metadata service
 - ISO: translate catalogs into ISO documents

(*) UDDC former acronym for ACDD

3.2 Enabling ncISO

- ThreddsConfig.xml:

```
<NCISO>  
  <ncmlAllow>true</ncmlAllow>  
  <uddcAllow>true</uddcAllow>  
  <isoAllow>true</isoAllow>  
</NCISO>
```

- ncISO for datasets in your catalogs:

```
<service name="ncml" serviceType="NCML" base="/thredds/ncml/" />  
<service name="uddc" serviceType="UDDC" base="/thredds/uddc/" />  
<service name="iso" serviceType="ISO" base="/thredds/iso/" />
```

3.3 nclSO exercise

1. Enable nclSO
2. Check your UDDC(*) score
3. Add metadata and improve the UDDC score

(*) UDDC former acronym for ACDD