

OPeNDAP, OpenSearch, & the CMR

Overview

One of the items on the BEDI list to increase data access involved exposing more products via OPeNDAP. We are working on documenting current usage of these access URLs in Discovery Metadata and developing recommendations for exposing it in our search APIs (CMR and OpenSearch flavors) and interpreting in our web clients ([Earthdata Search](#), [GCMD](#), [Reverb](#))

This is the first attempt at documenting the needs, the existing usage and recommendations, we want your feedback, so please feel free to comment below.

OpenSearch

Exposing OPeNDAP links in OpenSearch

Currently, CMR OpenSearch has no provision for exposing OPeNDAP links to the user.

We have been asked by GSFC to expose this information



WCIC-205 - Jira project doesn't exist or you don't have permission to view it.

SDPS are working on an OPeNDAP solution with NCR 8052122

In order to facilitate end user access to data it would be convenient to export OpenDAP URLs to ECHO so that the end user would have a choice of download not only the whole granule, but an OpenDAP specified subset of it. We plan to deploy OpenDAP at the DAACs before 8.3.3, so it would be useful if we could give the DAACs the ability to export OpenDAP URLs to ECHO so they could make use of it.

Drivers

1. CMR OpenSearch should present OPeNDAP urls compliant to ESIP Best Practices (see below)
2. CMR search results in ATOM format need to facilitate that
3. Providers need to supply enough information to do (1) and (2)
 - a. SDPS
 - b. Others
4. Other formats?

Discussion topics

1. How are SDPS exposing OPeNDAP url information to ECHO. Answer: SDPS are able to present the following, utilizing the OnlineResourceURL element -> url + type=OPeNDAP + mime-type of the originating resource (probably HDF)
2. ~~Is the collection and granule schema able to support the three elements required to comply with ESIP Best Practices?~~ Not unless we export a url for each explicit format.
3. How do we expose that in search for
 - a. ECHO10 - OnlineResourceURL
 - b. ISO - unknown
 - c. ATOM - link element with some discriminator saying 'this is an OPeNDAP url'
 - d. ATOM OpenSearch - see Best Practices

The ESIP Best Practice

The [ESIP OpenSearch Best Practices](#) documents a means to expose these URLs in a meaningful way. It is detailed below:

7.2.7 Characterizing OPeNDAP Link Elements (Optional)

- [Overview](#)
- [OpenSearch](#)
 - [Exposing OPeNDAP links in OpenSearch](#)
 - [Drivers](#)
 - [Discussion topics](#)
 - [The ESIP Best Practice](#)
 - [7.2.7 Characterizing OPeNDAP Link Elements \(Optional\)](#)
 - [OpenSearch and CMR Metadata](#)
- [CMR](#)
 - [Metadata URL Support](#)
 - [UMM-C](#)
 - [ECHO](#)
 - [DIF 9.9](#)
 - [Current Usage](#)
 - [Granule](#)
 - [Collection](#)
 - [Recommendations](#)
 - [Granule](#)
 - [Collection](#)
- [Additional Resources](#)

In addition to these basic link types, some data providers may offer a variety of data access points to the same data. An example of this is data providers that offer data via the OPeNDAP standard, which itself has several different return types: specific forms of metadata or data in varying formats. In order to distinguish OPeNDAP data access points from simple FTP or HTTP URLs, as well as from each other, an xlink-based scheme has been adapted to identify the type of OPeNDAP access point. An OPeNDAP service link is identified by a <link> attribute of the type "xlink:role". In this case, specifically, the xlink role and type attributes are given as:

xlink:role="http://xml.opendap.org/dap/dap2.xsd" xlink:type="simple"

In addition, the combination of the rel attribute, mime-type (i.e., type) attribute and xlink:arcrole attribute are used to distinguish the different types of OPeNDAP response:

Server Response Type	rel	type	xlink:arcrole
HTML	describedby	text/html	#html
info	describedby	text/html	#info
DDX	describedby	application/xml	#ddx
RDF	describedby	application/rdf+xml	#rdf
ASCII	enclosure	text/plain	#ascii
NetCDF	enclosure	application/x-netcdf	#netcdf

Note that a search engine is not required to return all of these responses when identifying OPeNDAP links, just the ones that it (a) supports and (b) wishes to advertise to clients.

It is understood that inclusion of these xlink elements can cause a Discovery response to fail validation by some Atom feed validators, but this has been traded off for the extra information this construct provides to client programs about the options for obtaining data and metadata from the provider.

The drivers behind this, at first glance, verbose means of describing an OPeNDAP service are as follows:

- **OpenSearch wishes to expose all possible formats to a client that doesn't understand OPeNDAP.**

An OPeNDAP-savvy client could take a URL to a resource and, knowing that it was an OPeNDAP URL, infer that appending various extension would yield meaningful results.

A client that doesn't understand OPeNDAP would not be able to exploit a net-cdf version of this resource without being explicitly told that.

So ESIP recommends that all formats of the resource be explicitly called out. For example,

the OPeNDAP resource

```
http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz
```

Would be represented in an OpenSearch result as:

```

<entry>
...
<link href="http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.html"
  hreflang="en-US" title="(OPeNDAP)" rel="describedby" type='text/html'
  xmlns:xlink="http://www.w3.org/1999/xlink" xlink:role="http://xml.opendap.org/dap/dap2.xsd"
  xlink:type="simple" xlink:arcrole='#html' />
<link href="http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.info"
  hreflang="en-US" title="(OPeNDAP)" rel="describedby" type='text/html'
  xmlns:xlink="http://www.w3.org/1999/xlink" xlink:role="http://xml.opendap.org/dap/dap2.xsd"
  xlink:type="simple" xlink:arcrole='#info' />
<link href="http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.ddx"
  hreflang="en-US" title="(OPeNDAP)" rel="describedby" type='application/xml'
  xmlns:xlink="http://www.w3.org/1999/xlink" xlink:role="http://xml.opendap.org/dap/dap2.xsd"
  xlink:type="simple" xlink:arcrole='#ddx' />
<link href="http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.rdf"
  hreflang="en-US" title="(OPeNDAP)" rel="describedby" type='application/rdf+xml'
  xmlns:xlink="http://www.w3.org/1999/xlink" xlink:role="http://xml.opendap.org/dap/dap2.xsd"
  xlink:type="simple" xlink:arcrole='#rdf' />
<link href="http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.cdf"
  hreflang="en-US" title="(OPeNDAP)" rel="enclosure" type='application/x-netcdf'
  xmlns:xlink="http://www.w3.org/1999/xlink" xlink:role="http://xml.opendap.org/dap/dap2.xsd"
  xlink:type="simple" xlink:arcrole='#netcdf' />
<link href="http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.ascii"
  hreflang="en-US" title="(OPeNDAP)" rel="enclosure" type='text/plain'
  xmlns:xlink="http://www.w3.org/1999/xlink" xlink:role="http://xml.opendap.org/dap/dap2.xsd"
  xlink:type="simple" xlink:arcrole='#ascii' />
...
</entry>

```

- **OpenSearch wishes to convey, explicitly, the capabilities of an OPeNDAP service since not all versions of OPeNDAP support all formats**

Meaning the above result represents an up-to-date OPeNDAP server. Other servers may not support net-cdf fo

OpenSearch and CMR Metadata

CMR OpenSearch decorates a CMR ATOM response, at the collection and granule levels, with OpenSearch specific elements and augmentations.

In order to expose OPeNDAP url information as per the ESIP recommendations we need to know the following,

- that a resource falls in the OPeNDAP category
- What mime type it is
- what subset, if any, of OPeNDAP formats it supports

Currently, the ECHO10 result format will identify a url of type 'OPeNDAP' as follows,

```

curl https://cmr.earthdata.nasa.gov:443/search/concepts/G550016-GHRC

<OnlineResource>
  <URL>http://ghrc.nsstc.nasa.gov/opendap/ssmi/f14/monthly/</URL>
  <Type>OPeNDAP</Type>
</OnlineResource>

```

That same resource in ATOM gives (note CMR is still using the 'old' relations to describe a link)

```

curl https://cmr.earthdata.nasa.gov:443/search/concepts/G550016-GHRC.atom

<link href="http://ghrc.nsstc.nasa.gov/opendap/ssmi/f14/monthly/" hreflang="en-US" title="(OPeNDAP)" rel="
http://esipfed.org/ns/fedsearch/1.1/data#"></link>

```

In order for CMR OpenSearch to expose this as an OPeNDAP resource we would need from CMR, at a minimum, some means of identifying it as an OPeNDAP url.

If we wanted to convey what formats that OPeNDAP implementation supports we have a bigger problem. That information would need to be obtained from the provider.

CMR

Metadata URL Support

UMM-C

In the Unified Metadata Model for Collections, all URLs are mapped into the RelatedURL element. This element looks almost identical to the existing DIF Related_URL.

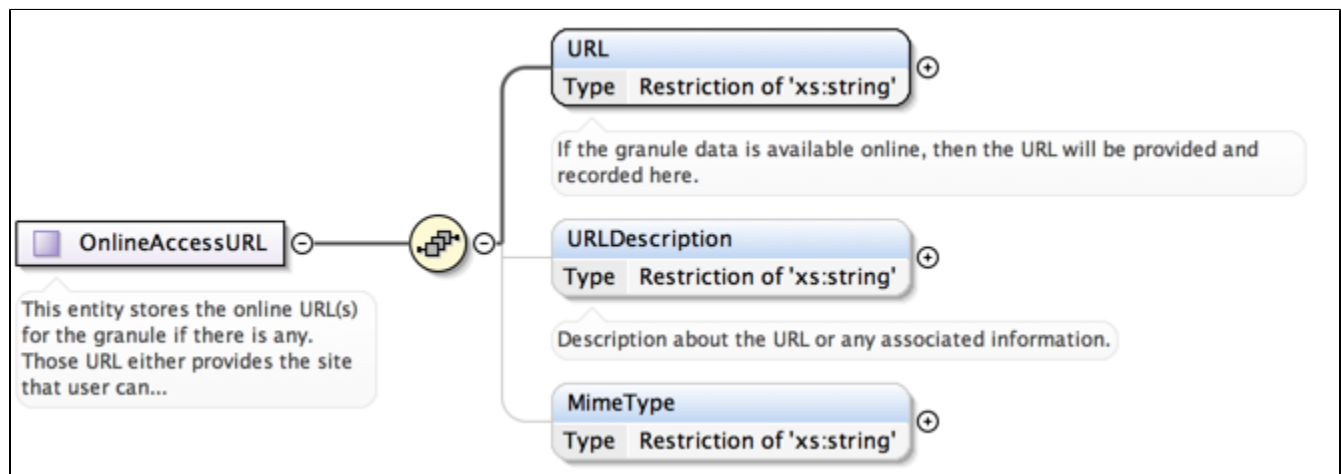
Here is a sample mapping from UMM-C to the DIF and ECHO 10 elements:

UMM-C Element Specification	DIF Mapping	ECHO 10 Online Access	ECHO 10 Online Resource
RelatedUrl/ContentType/Type	RelatedUrl/ContentType/Type	-	OnlineResource/Type*
RelatedUrl/ContentType/Subtype	RelatedUrl/ContentType/Subtype	-	OnlineResource/Type*
RelatedUrl/URL	RelatedUrl/URL	OnlineAccessURL/URL	OnlineResource/URL
RelatedUrl/Description	RelatedUrl/Description	OnlineAccessURL/Description	OnlineResource/Description
RelatedUrl/MimeType	RelatedUrl/MimeType	OnlineAccessURL/MimeType	OnlineResource/MimeType

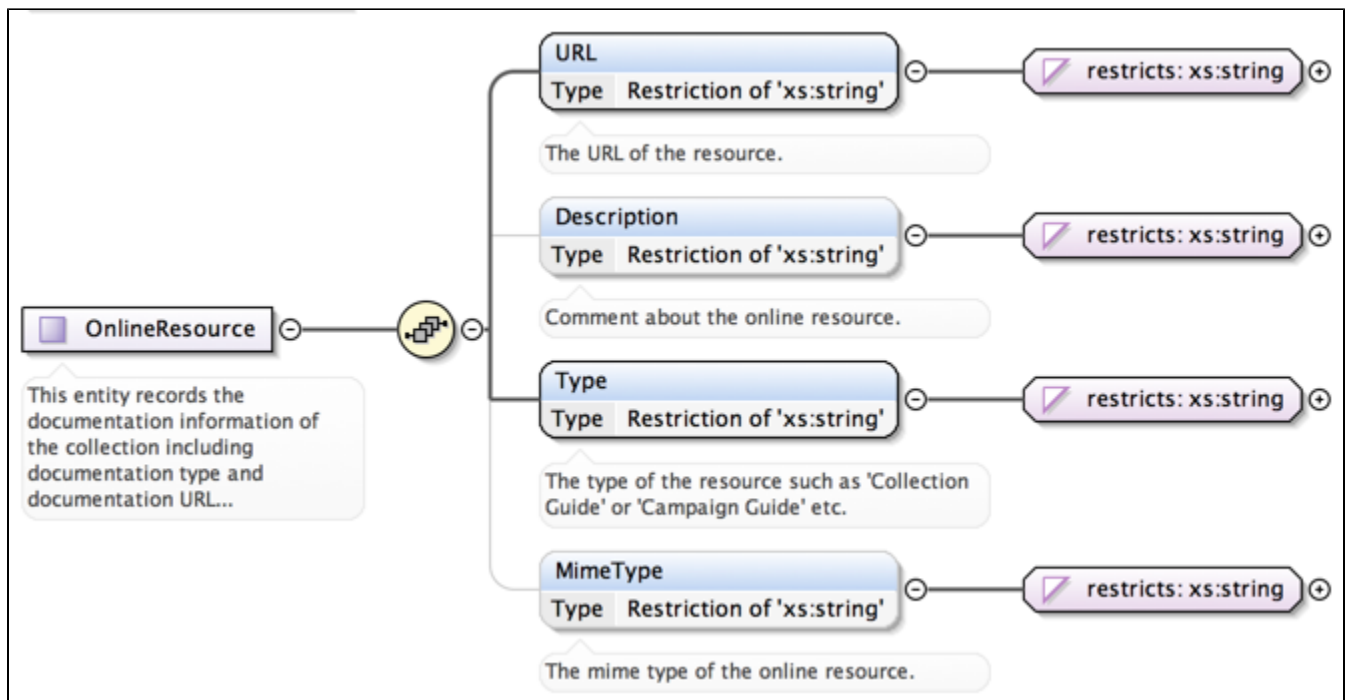
* The ECHO OnlineResource/Type element is used by some DAACs to provide Subtype via a colon delimited string (e.g. 'GET DATA : OPENDAP DATA (DODS)').

ECHO

via Online Access URL

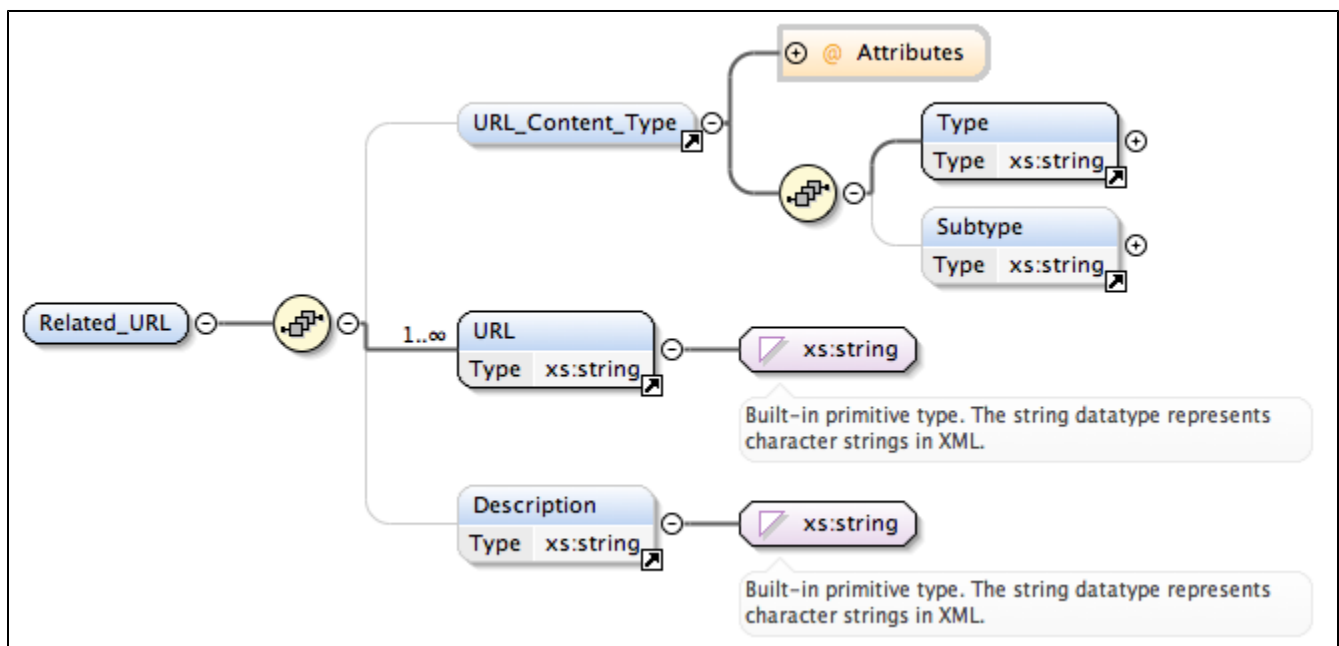


via Online Resource



DIF 9.9

via Related_URL



Current Usage

Granule

ECHO 10 Format

GHRC Implementaion

```
(https://cmr.earthdata.nasa.gov:443/search/concepts/G550016-GHRC)
<OnlineResource>
  <URL>http://ghrc.nsstc.nasa.gov/opendap/ssmi/f14/monthly/</URL>
  <Type>OPeNDAP</Type>
</OnlineResource>
```

PO.DAAC Implementaion

```
(https://cmr.earthdata.nasa.gov:443/search/concepts/G2921045-PODAAC)
<OnlineAccessURL>
  <URL>http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly/
osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.html</URL>
  <URLDescription>The OPeNDAP location for the granule.</URLDescription>
</OnlineAccessURL>
```

SDPS Implementaion

```
(http://testbed.echo.nasa.gov/reverb/granules/G1000021178-EDF_OPS)
<OnlineAccessURL>
  <URL>http://f5eil01v.edn.ecs.nasa.gov:24336/opendap//FS1/MOST/MOD29P1D.005/2011.01.01/MOD29P1D.A2011001.
h13v27.005.2011002082448.hdf.html</URL>
  <MimeType>text/html</MimeType>
</OnlineAccessURL>
```

Collection

ECHO 10 Format

GES DISC Implementaion (Collection Level)

```
(https://cmr.earthdata.nasa.gov:443/search/concepts/C186901238-GSFCS4PA)
<AdditionalAttributes>
  <AdditionalAttribute>
    <Name>OPeNDAPServer</Name>
    <DataType>STRING</DataType>
    <Description>Access the AIRS L2 Cloud-Cleared Infrared (IR) Radiances Product through OPeNDAP.<
/Description>
    <Value>http://airs12.gesdisc.eosdis.nasa.gov/opendap/Aqua_AIRS_Level2/AIRI2CCF.005/contents.html</Value>
  </AdditionalAttribute>
</AdditionalAttributes>
```

Alternate GES DISC Implementaion (Collection Level)

```
(https://cmr.earthdata.nasa.gov:443/search/concepts/C114367117-GSFCS4PA)
<OnlineResource>
  <URL>http://disc2.nascom.nasa.gov/opendap/TRMM_L3/TRMM_CSH/</URL>
  <Description>OPeNDAP Access dataset TRMM Monthly 0.5 degree x 0.5 degree Convective/Stratiform Heating<
/Description>
  <Type>GET DATA : OPeNDAP DATA (DODS)</Type>
</OnlineResource>
```

DIF Format

GES DISC Implementation

```
(http://gcmd.gsfc.nasa.gov/KeywordSearch/Metadata.do?
Portal=GCMD&MetadataType=0&MetadataView=Full&KeywordPath=&EntryId=GES_DISC_TRMM_CSH_V6)
<Related_URL>
  <URL_Content_Type uuid="2f1f3f44-61d5-48db-a8ac-06e3f777bf35">
    <Type>GET DATA</Type>
    <Subtype>OPENDAP DATA (DODS)</Subtype>
  </URL_Content_Type>
  <URL>http://disc2.nascom.nasa.gov/opendap/TRMM_L3/TRMM_CSH/</URL>
  <Description>
    OPENDAP Access dataset TRMM Monthly 0.5 degree x 0.5 degree Convective/Stratiform Heating
  </Description>
</Related_URL>
```

Recommendations

Granule

ECHO 10 Format

Providers will describe their OPeNDAP urls, at a granule level, as follows

GHRC Implementaion

```
<OnlineResource>
  <URL>http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly
/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz</URL>
  <Type>GET DATA : OPENDAP DATA (DODS)</Type>
  <MimeType>application/x-gzip</MimeType>
</OnlineResource>
```

Advantages

Given this information a client could, understanding it was an OPeNDAP resource, invoke services on that granule by changing the extension and/or query string to the url. For example,

- GET http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.html - to get a form for processing the file
- GET http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.ascii - to get an ascii representation of the data
- GET http://podaac-opendap.jpl.nasa.gov/opendap/allData/coastal_alt/preview/L4/OSU_COAS/weekly/osu_cioss_weekly_msla_geovel_1992_v1.nc.gz.ddx - to get a description of the capabilities of the OPeNDAP resource
- netcdf?

Collection

ECHO 10 Format

Providers may describe collections of granules with the same OPeNDAP capabilities at the collection level as follows,

GES DISC Implementaion (Collection Level)

```
<OnlineResource>
  <URL>http://airs12.gesdisc.eosdis.nasa.gov/opendap/Aqua_AIRS_Level12/AIRI2CCF.005/contents</URL>
  <Type>GET DATA : OPENDAP DATA (DODS)</Type>
</OnlineResource>
```

DIF Format

Current usage by GES DISC is good. We recommend sticking with this.

GES DISC Implementation

```
(http://gcmd.gsfc.nasa.gov/KeywordSearch/Metadata.do?
Portal=GCMD&MetadataType=0&MetadataView=Full&KeywordPath=&EntryId=GES_DISC_TRMM_CSH_V6)
<Related_URL>
  <URL_Content_Type uuid="2f1f3f44-61d5-48db-a8ac-06e3f777bf35">
    <Type>GET DATA</Type>
    <Subtype>OPeNDAP DATA (DODS)</Subtype>
  </URL_Content_Type>
  <URL>http://disc2.nascom.nasa.gov/opendap/TRMM_L3/TRMM_CSH/</URL>
  <Description>
    OPeNDAP Access dataset TRMM Monthly 0.5 degree x 0.5 degree Convective/Stratiform Heating
  </Description>
</Related_URL>
```

Advantages

Given this information a client could, understanding it was an OPeNDAP resource, invoke services on that granule by changing the extension and/or query string to the url. For example,

- GET http://airs12.gesdisc.eosdis.nasa.gov/opendap/Aqua_AIRS_Level2/AIRI2CCF.005/contents.ddx - to get a description of the capabilities of the OPeNDAP resources that are children of this resource

Using this technique a client could infer the properties of many granule resources at a time without having to grab a ddx file for each granule.

Additional Resources

[Setting up a local OPeNDAP server with Docker](#)