GCMD OpenSearch

GCMD OpenSearch

Existing Code

Tag in all pages, search results are portal (metadata subsets) specific:

```
<link rel="search" type="application/opensearchdescription+xml" title="Global Change Master Directory"
href="/KeywordSearch/default/openSearch.jsp?Portal=GCMD">
```

Which produces the following:

```
<OpenSearchDescription xmlns="http://a9.com/-/spec/opensearch/1.1/">
    <Tags>
        agriculture atmosphere biosphere cryosphere snow ice climate hydrosphere land ocean paleoclimate
sun earth services
    </Tags>
    <ShortName>GCMD</ShortName>`
    <Image height="16" width="16" type="image/x-icon">http://gcmd.nasa.gov/favicon.ico</Image>
    <Contact>gsfc-gcmduso@mail.nasa.gov</Contact>
    <LongName>Global Change Master Directory</LongName>
    <Description>
        The GCMD database holds more than 25,000 descriptions of Earth science data sets and services
covering all aspects of Earth and environmental sciences.
    </Description>
    <Query role="example" searchTerms="modis"/>
    <Developer>GCMD Development Team</Developer>
    <Attribution>Search data copyright GCMD/NASA</Attribution>
    <SyndicationRight>limited</SyndicationRight>
    <InputEncoding>UTF-8</InputEncoding>

    <Url type="text/html" method="GET"
template="http://gcmd.gsfc.nasa.gov:80/KeywordSearch/OpenSearch.do?searchTerms={searchTerms}&output=html
&count={count}&startIndex={startIndex}&startPage={startPage}&MetadataType=0&Portal=GCMD"/>

    <Url type="application/rss+xml" method="GET"
template="http://gcmd.gsfc.nasa.gov:80/KeywordSearch/OpenSearch.do?searchTerms={searchTerms}&output=rss&
count={count}&startIndex={startIndex}&startPage={startPage}&geoBox={geo:box}&timeStart={time:start}&time
End={time:end}&MetadataType=0&Portal=GCMD"/>

    <Url type="application/atom+xml" method="GET"
template="http://gcmd.gsfc.nasa.gov:80/KeywordSearch/OpenSearch.do?searchTerms={searchTerms}&output=atom
&count={count}&startIndex={startIndex}&startPage={startPage}&geoBox={geo:box}&timeStart={time:start}&tim
eEnd={time:end}&MetadataType=0&Portal=GCMD"/>

</OpenSearchDescription>
```

From this we can see that GCMD supports:

- Searches
  - terms - `{searchTerms}`
  - pagination - `{count}`, `{startIndex}`, `{startPage}`
  - geo - `{geo:box}` (code assumes `INTERSECTS` mode)
  - date - `{time:start}`, `{time:end}` (code assumes `T00:00:00Z`)
- Formats
  - HTML
  - RSS
  - ATOM
  - CSV (New - not documented yet)

Examples:
http://gcmd.gsfc.nasa.gov/KeywordSearch/OpenSearch.do?searchTerms=water&output=rss&count=10&startIndex=1&startPage=1&Portal=GCMD

http://gcmd.gsfc.nasa.gov/KeywordSearch/OpenSearch.do?searchTerms=water&output=rss&count=3&startIndex=1&startPage=1&geoBox=-150,-45,-150,45&Portal=GCMD

http://gcmd.gsfc.nasa.gov/KeywordSearch/OpenSearch.do?searchTerms=water&output=rss&count=3&startIndex=1&startPage=1&geoBox=-150,-45,-150,45&timeStart=2000-02-14&timeEnd=2010-02-14&Portal=GCMD

Output from the last is:

```
<rss version="2.0"
        xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/"
        xmlns:atom="http://www.w3.org/2005/Atom"
        xmlns:georss="http://www.georss.org/georss"
        xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/"
        xmlns:time="http://a9.com/-/opensearch/extensions/time/1.0/">
    <channel>
        <title>GCMD Search: [Freetext=&#039;water&#039;]|[Spatial_Coverage
INTERSECTS:Northernmost_Latitude=&#039;45&#039;,Southernmost_Latitude=&#039;-45&#039;,Easternmost_Longit
ude=&#039;-150&#039;,Westernmost_Longitude=&#039;-150&#039;]|[Temporal_Coverage:
Start_Date=&#039;2000-02-14&#039;, Stop_Date=&#039;2010-02-14&#039;] </title>

<link>http://gcmd.gsfc.nasa.gov/KeywordSearch/OpenSearch.do?searchTerms=water&amp;output=html&amp;Metada
taType=0&amp;count=1&amp;startIndex=1&amp;startPage=1&amp;geoBox=-150,-45,-150,45&amp;timeStart=2000-02-
14T00:00:00Z&amp;timeEnd=2010-02-14T00:00:00&amp;Portal=GCMD</link>

        <description>Search results for [Freetext=&#039;water&#039;]|[Spatial_Coverage
INTERSECTS:Northernmost_Latitude=&#039;45&#039;,Southernmost_Latitude=&#039;-45&#039;,Easternmost_Longit
ude=&#039;-150&#039;,Westernmost_Longitude=&#039;-150&#039;]|[Temporal_Coverage:
Start_Date=&#039;2000-02-14&#039;, Stop_Date=&#039;2010-02-14&#039;]</description>

        <opensearch:totalResults> 1193 </opensearch:totalResults>
        <opensearch:startIndex> 0 </opensearch:startIndex>
        <opensearch:itemsPerPage>1</opensearch:itemsPerPage>

        <atom:link rel="search" type="application/opensearchdescription+xml"
href="http://gcmd.gsfc.nasa.gov/KeywordSearch/default/opensearchdescription.xml"/>

        <item>
            <title>NOAA/GFDL IPCC Special Report on Emission Scenarios 720 ppm Stabilization Experiment
(SRES A1BA2) Model Run 1, CM2.1</title>
            <link>http://gcmd.gsfc.nasa.gov/getdif.htm?[GCMD]NOAA_GFDL_CM2.1U-H2_SresA1B_X1</link>
            <description>
                GFDL experiment name = CM2.1U-H2_SresA1B_X1. PCMDI experiment name =
SRESA1B (run1). Initial conditions for this experiment were taken from
1 January 2001 of the climate of the 20th century experiment ...
            </description>
        </item>

    </channel>
</rss>
```

An undocumented feature, CSV output:

```
curl -i
"http://gcmd.gsfc.nasa.gov/KeywordSearch/OpenSearch.do?searchTerms=water&output=csv&count=10&startIndex=
10&startPage=1&geoBox=-150,-45,-150,45&timeStart=2000-02-14T00:00:00Z&timeEnd=2010-02-14T00:00:00&Portal
=GCMD"
```

Which produces:

```
row,    Orig_Metadata_ID,   Entry_ID,   Parent_EntryId, Has Children, Last_Revision_Date
1,  GCMD,   SHADOZ_TAHITI,  SHADOZ ,    no, Tue Jun 09 00:00:00 EDT 2009
2,  GCMD,   CEADO_Physical_Chemical,    no parent, no, Tue Sep 27 00:00:00 EDT 2005
3,  GCMD,   AIRS_EPA,   no parent,  no, Tue May 08 00:00:00 EDT 2012
4,  GCMD,   OSDPD0007_08,   no parent,  no, Wed Oct 01 00:00:00 EDT 2003
5,  GCMD,   OSDPD0014_16_20,    no parent,  no, Wed Sep 25 00:00:00 EDT 1996
6,  GCMD,   USGS_GEOCHEM,   no parent,  no, Tue Jan 25 00:00:00 EST 2005
7,  GCMD,   OAR0006,    no parent,  no, Thu Oct 26 00:00:00 EDT 2006
8,  GCMD,   OAR0007,    no parent,  no, Wed Aug 31 00:00:00 EDT 1994
9,  GCMD,   FW00061,    no parent,  no, Mon Feb 28 00:00:00 EST 2005
10, GCMD,   FW00064,    no parent,  no, Wed Aug 07 00:00:00 EDT 2002
```

Future

For metadata with a Related URL of type `OpenSearch-Granule` (example), include those urls in the output of the atom/rss feed for each item to allow for the chaining of OpenSearch calls. Want to talk with outers about how to denote the OpenSearch URLs as different from any URL GCMD may give to point to our metadata or do we add a flag to our interface to state that the feeds are to include opensearch result URLs over the GCMD `Get Data` link

RSS sudo code:

```
<rss><item><link>http://echo.gsfc.nasa.gov/blablabla</link></item></rss>
```

Atom sudo code:

```
<atom><entry><link href="http://echo.gsfc.nasa.gov/blablabla"></link></rss>
```

Did find some attributes for atom link tag[1]:

```
<xs:complexType name="linkType" mixed="true">
    <xs:annotation>
        <xs:documentation>
            The Atom link construct is defined in section 3.4 of the format spec.
        </xs:documentation>
    </xs:annotation>
    <xs:attribute name="href" use="required" type="xs:anyURI"/>
    <xs:attribute name="rel" type="xs:string" use="optional"/>
    <xs:attribute name="type" use="optional" type="xs:string"/>
    <xs:attribute name="hreflang" use="optional" type="xs:NMTOKEN"/>
    <xs:attribute name="title" use="optional" type="xs:string"/>
    <xs:attribute name="length" use="optional" type="xs:positiveInteger"/>
    <xs:attributeGroup ref="atom:commonAttributes"/>
</xs:complexType>
```

---

1. http://www.kbcafe.com/rss/atom.xsd.xml