Accessing Metadata with GCMD Interfaces

Metadata Web Service (MWS)
Keyword Management Service (KMS)
OpenSearch and RSS via Web

thomas.a.cherry@nasa.gov

Thu Jan 09 2014 - 37 slides
Common KMS/MWS Architecture

GET, POST, HEAD

Access Manager

RESTful API Layer (KMS & MWS)

Data API Layer

Relational Database

Cache

web

ESDIS Authentication DB

GCMD Authorization DB
Metadata Web Service (MWS)

Keyword Management Service (KMS)

OpenSearch and RSS via Web
• Legacy Interfaces
  • Replaced OpenAPI
  • Designed to be binary compatible with OpenAPI
  • OpenAPI around a little longer...
• New Interfaces such as ticket numbers
• Internally Supports:
  • Document Builder
  • Bulk Uploader
MWS Flow

http://gcmd.nasa.gov/Connect/ (HTML)

http://gcmdservices.gsfc.nasa.gov/mws/ (API) (Last slash is significant)

GET DIF/SERF/CD
GET multiple DIFs/SERFs/CDs (limitations, cached)
GET Entry IDs (query)
GET Keywords (valids)
GET Supplementals
POST Publisher
ticket number
GET status (not listed)
POST Validator...
GET status (not listed)
## MWS Examples

<table>
<thead>
<tr>
<th>Language</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>cURL</td>
<td>$(curl -u demo:friend &quot;<a href="http://gcmdservices.gsfc.nasa.gov/mws/">http://gcmdservices.gsfc.nasa.gov/mws/</a>&quot;)</td>
</tr>
<tr>
<td>Java</td>
<td><a href="http://gcmd.nasa.gov/Connectclients/mws/SimpleJavaClient.java">http://gcmd.nasa.gov/Connectclients/mws/SimpleJavaClient.java</a></td>
</tr>
<tr>
<td>Objective-c</td>
<td>command line app objective-c in xcode</td>
</tr>
</tbody>
</table>
Metadata Web Service (MWS)

Keyword Management Service (KMS)
OpenSearch and RSS via Web
KMS Overview

- Active Development (Related Keywords)
- SKOS Data Model
- Used by GCMD to Support:
  - Keyword Manager
  - /static/ views
  - Web Interface
SKOS - Simple Knowledge Organization System (See W3C)

Example scheme:
- sciencekeywords
- platforms
- projects

- Only one broader relationship per concept, same scheme
- Multiple narrower relationships, same scheme
- Concepts have alt labels
- Related concept relationships which cross schemes (maybe, future)
<conceptBrief id="19871" uuid="1e005a-312c-4d74-8d42-6f1ad758f999" prefLabel="Science Keywords" conceptSchemeId="31" conceptScheme="sciencekeywords" isLeaf="false" status=""/>

<conceptBrief id="23367" uuid="713eb469-a6e4-446b-8b6-1341b7e9d8b8" prefLabel="Locations" conceptSchemeId="33" conceptScheme="locations" isLeaf="false" status=""/>

<conceptBrief id="24337" uuid="a15bf804-b000-4013-831e790381c1" prefLabel="Producers" conceptSchemeId="35" conceptScheme="producers" isLeaf="false" status=""/>

<conceptBrief id="31173" uuid="e3261b5d-34c1-4980-af2f-f97e7e2b6d12" prefLabel="Platforms" conceptSchemeId="37" conceptScheme="platforms" isLeaf="false" status=""/>

<conceptBrief id="32591" uuid="b2140599-b0a7-415c-9513-cee79783ff8" prefLabel="Instruments" conceptSchemeId="39" conceptScheme="instruments" isLeaf="false" status=""/>

<conceptBrief id="35209" uuid="60b9fed-5c96-4989-8c64-a479bbed3ab" prefLabel="Projects" conceptSchemeId="41" conceptScheme="projects" isLeaf="false" status=""/>

<conceptBrief id="39447" uuid="875a63-ac04-4136-bc25-0c008ec00b95" prefLabel="Related URL Content Types" conceptSchemeId="49" conceptScheme="rucontenttype" isLeaf="false" status=""/>

<conceptBrief id="39531" uuid="1499785c-8b74-45f4-bbf7-192d4dd43b2f" prefLabel="Horizontal Resolution Ranges" conceptSchemeId="51" conceptScheme="horizontalresolutionrange" isLeaf="false" status=""/>

<conceptBrief id="39563" uuid="972ca44-7b8-4a4-a9d2-49f0a47af5" prefLabel="Vertical Resolution Ranges" conceptSchemeId="53" conceptScheme="verticalresolutionrange" isLeaf="false" status=""/>

<conceptBrief id="39579" uuid="502ad03c-db64-4b32-8aa-13fa97c3988" prefLabel="Temporal Resolution Ranges" conceptSchemeId="55" conceptScheme="temporalresolutionrange" isLeaf="false" status=""/>
Upcoming KMS features

Almost all interfaces are cached

HTTP Headers/HEAD Requests:
- Content-MD5: 32875a669eade69fc366679c97905791

Extensions:
- `curl -I -u esip:esip2014` 
  * `http://gcmdservices.gsfc.nasa.gov/kms/concept/c5a4126a-8f0a-4774-9bef-c2a9f9487727`
- `curl -I -u esip:esip2014` 
  * `http://gcmdservices.gsfc.nasa.gov/kms/concept/c5a4126a-8f0a-4774-9bef-c2a9f9487727.rdf`
- `curl -I -u esip:esip2014` 
  * `http://gcmdservices.gsfc.nasa.gov/kms/concept/c5a4126a-8f0a-4774-9bef-c2a9f9487727.xml`
### KMS Examples

<table>
<thead>
<tr>
<th>Language</th>
<th>Code Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>cURL</td>
<td><code>curl -u esip:esip2014 &quot;http://gcmdservices.gsfc.nasa.gov/kms/&quot;</code></td>
</tr>
<tr>
<td>Java</td>
<td><code>http://gcmd.nasa.gov/Connectclients/kms/SimpleJavaClient.java</code></td>
</tr>
<tr>
<td>Objective-c</td>
<td><code>app objective-c in xcode</code></td>
</tr>
</tbody>
</table>
• Created to provide unauthenticated access
• Layout similar to KMS RESTful urls
• Exported from KMS every 6 hours, exact same data
• Multiple formats (CSV, OWL, RDF, XML)
• Compressed files available for some concept files.
• One historic copy maintained for most concepts
## Static Interface vs KMS

<table>
<thead>
<tr>
<th>Pro</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No Authentication</td>
<td>• URLs reference KMS API...</td>
</tr>
<tr>
<td>• Fast</td>
<td>• No search functionality</td>
</tr>
<tr>
<td>• Compressed files and hashes</td>
<td>• No custom HTTP Header Tags</td>
</tr>
</tbody>
</table>
Metadata Web Service (MWS)
Keyword Management Service (KMS)
OpenSearch and RSS via Web
The software behind the GCMD web site is not a "service" like MWS or KMS, but it does have some really service like interesting features

http://gcmd.gsfc.nasa.gov/KeywordSearch/
- OpenSearch
- RSS feed for all search results
- GCMD RSS/ATOM feeds
RDF/ATOM Feeds

- RSS feed for all search results
- GCMD RSS/ATOM feeds
  - Next release, new GCMD feed backend
  - New Records, Modified Records
  - Deleted Records (new)
• OpenSearch
  • Initially only browser integration
  • Distributed searching frequently requested
  • Documentation to be added to the /Connect/ page

http://gcmd.nasa.gov/Connect/
GCMD supports:

- Searches
  - terms - {searchTerms}
  - pagination - {count}, {startIndex}, {startPage} (optional)
  - geo - {geo:box} (code assumes INTERSECTS mode) (optional)
  - date - {time:start}, {time:end} (code assumes T00:00:00Z) (optional)

- Formats
  - HTML
  - RSS
  - ATOM
  - CSV (New - not documented, columns may change)
  - Plans for KML
http://gcmd.nasa.gov/r/u/report