



CEOS OpenSearch
Best Practice Document
Version 1.0

CEOS Document

[CEOS-OPENSEARCH-BP-V001]

Revision History

Date	Ver	Changes
2014-12-17	1.0(draft)	Draft release for public comment period
2015-05-29	1.0	Official release

Table of Contents

1 Introduction

1.1 Purpose of the document

1.2 Project Member

1.3 Terminology

Terminology referring to data granularity

Granule

Collection

CWIC (CEOS WGISS Integrated Catalog)

CEOS IDN (CEOS International Directory Network)

FedEO (Federated Earth Observation Missions)

OGC (Open Geospatial Consortium)

1.4 Specifications and references

(1) OpenSearch specifications and extensions

OpenSearch Specification, 1.1, Draft 5

OpenSearch Geo and Time Extensions [OGC 10-032r8]

OpenSearch Extension for EO Products [OGC 13-026r5]

OpenSearch Parameter Extension - Draft 2

OpenSearch Relevance extension - Draft 1

O&M (Observation and Measurement) EO Profile [OGC 10-157r4]

(2) Reference

1.5 What is OpenSearch ?

Discovery

Description of Service

Search

Search Response

2 CEOS OpenSearch Best Practices

2.1. Two Step Search

CEOS-BP-001 - Support of two step search [Recommended]

CEOS-BP-002 - dc:identifier in collection search response [Requirement]

2.2 OSDD

CEOS-BP-003 - Support of OpenSearch Parameter Extension (Draft 2)

[Recommended]

Free Text Keyword

Geometry

CEOS-BP-004 - rel attribute of the URL in OSDD [Recommended]

CEOS-BP-005 - CEOS OpenSearch Best Practice identifier [Recommended]

2.3 Search Request

CEOS-BP-006 - Supported search parameters [Requirement]

- [CEOS-BP-007 - Multi-words for searchTerms \[Recommended\]](#)
- [CEOS-BP-008 - Use of startPage over startIndex \[Recommended\]](#)
- [CEOS-BP-009 - Search with geo:name \[Recommended\]](#)
- [CEOS-BP-010 - Output encoding format in search URL \[Optional\]](#)
- 2.4 [Search Response](#)
 - [CEOS-BP-011 - Output encoding format support \[Requirement\]](#)
 - [CEOS-BP-012 - Metadata representation in search response \[Recommended\]](#)
 - [CEOS-BP-013 - atom:summary \[Recommended\]](#)
 - [CEOS-BP-014 - GeoRSS \[Recommended\]](#)
 - [CEOS-BP-015 - Browse Image \[Recommended\]](#)
 - [CEOS-BP-016 - Data access \[Recommended\]](#)
- 2.5 [Exceptions](#)
- 2.6 [Future discussions](#)
- 3. [Closer look on implementations](#)
 - 3.1 [Basics](#)
 - [OSDD URL](#)
 - [Collection Level](#)
 - [Granule Level \(examples\)](#)
 - [Search URL](#)
 - [Collection Level](#)
 - [Granule Level \(examples\)](#)
 - [Documents](#)
 - 3.2 [Remarkable Practices](#)
 - [OSDD](#)
 - [Support of parameter extension](#)
 - [Support of Explain operation \(SRU\)](#)
 - [Output encoding format in a search URL](#)
 - [Support of 2 step search](#)
 - [Search Parameters](#)
 - [Supported Search Parameters](#)
 - [Supported output format](#)
 - [Output schema support](#)
 - [Free Keyword notation \(searchTerms\)](#)
 - [Pagination](#)
 - [Search Result](#)
 - [Browse](#)
 - [Thumbnail](#)
 - [Mask Image](#)
 - [Metadata](#)
 - [Documentation](#)
 - [Data](#)
 - [Error Handling](#)
 - [Exception Status Code \(4xx, 5xx\)](#)

[B1. OSDD \(Collection Level\)](#)

[IDN/CWIC](#)

[FedEO](#)

[CNES](#)

[B2. OSDD \(Granule Level\)](#)

[IDN/CWIC](#)

[FedEO](#)

[CNES](#)

[B3. Search request](#)

[IDN/CWIC](#)

[FedEO](#)

[CNES](#)

[B4. Search response \(Atom\)](#)

[IDN/CWIC](#)

[FedEO](#)

[CNES](#)

[2.5 Making A Catalog Search-Engine-Friendly](#)

1 Introduction

This document provides server implementation best practices of EO OpenSearch search service that allows for standardized and harmonized access to metadata and data of CEOS agencies, including CWIC and FedEO.

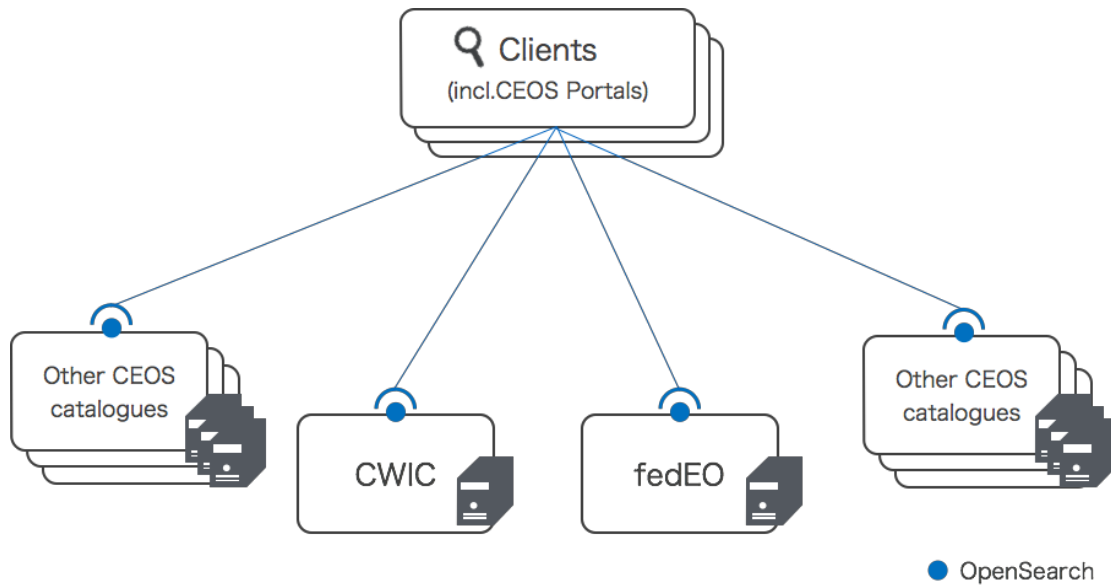


Figure 1 - CEOS OpenSearch

1.1 Purpose of the document

The purpose of this document is to achieve followings.

- Promote the use of the OpenSearch standard as a means of data discovery for Earth Data providers
- Define the expectations and requirements of candidate OpenSearch implementations
- Remove ambiguity in implementation where possible
- Facilitate the aggregation of results between disparate Earth Data providers via OpenSearch common standards
- Allow for clients to access search engines via an OpenSearch Description Document (OSDD) with no a priori knowledge of the interface
- Facilitate smooth integration between related OpenSearch implementations, such as a dataset resource collection that refers to granule resource collections from another provider

1.2 Project Member

This document has been developed by the CEOS OpenSearch Project Team of the Committee on Earth Observation Satellites (CEOS) Working Group on Information Systems and Services (WGISS).

The project team members are listed below.

JAXA	Yoshiyuki Kudo (co-lead of the project) Satoko Miura
CNES	Jérôme Gasperi (co-lead of the project) Richard Moreno
ESA	Pier Giorgio Marchetti Giuseppe Troina Yves Coene
NASA	Andrew Mitchell Yonsook Enloe Douglas Newman Christopher Lynnes
NOAA	Martin Yapur Kenneth McDonald
CCMEO	Patrick King

1.3 Terminology

Terminology referring to data granularity

Earth observation (EO) satellite data usually falls into two types in the context of granularity - collection and granule.

Granule

The finest granularity of data that can be independently managed. Granule usually matches the individual file of EO satellite data.

Collection

An aggregate of granules sharing the same product specification. A collection typically corresponds to the series of products derived from data acquired by a sensor on board a

satellite and having the same mode of operation.

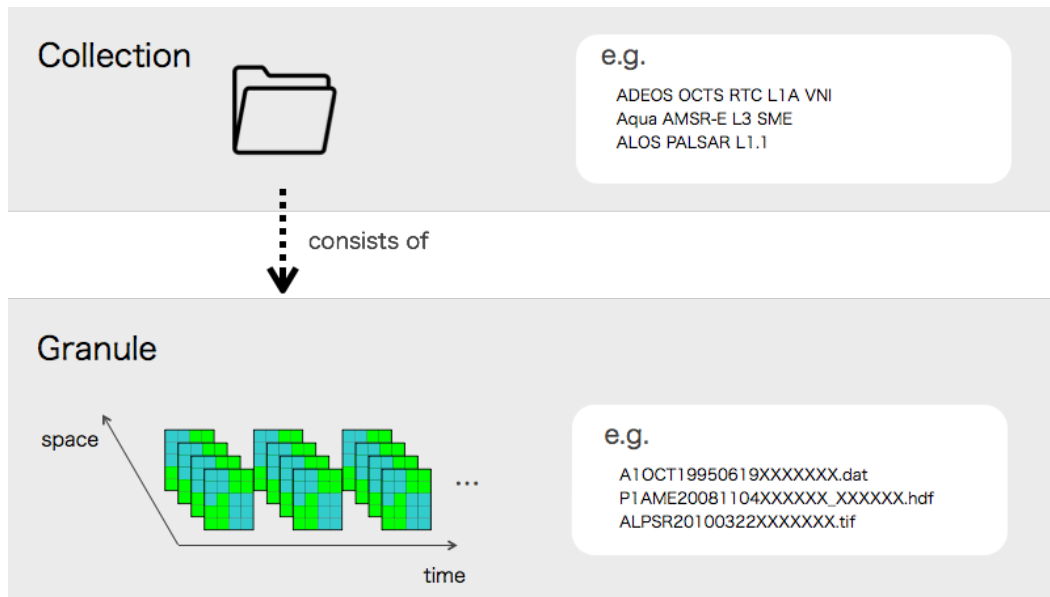


Figure 2 - EO satellite data granularity

Here, it is interesting to show how agencies use different terminology for referring to these two granularities.

Terminology (Fig 2)	ISO terms (i.e. ISO19115-2)	Agency dependent terminology
collection	dataset series	collection (CNES, NASA) dataset (JAXA) dataset series (ESA) product (JAXA)
granule	dataset	dataset (ESA) granule (NASA) product (ESA, CNES) scene (JAXA)

Since the agencies share the common notion of this two level hierarchy, the difference does not harm the harmonization effort of this activity.

The terminology “collection” and “granule” - are used throughout this document.

CWIC (CEOS WGISS Integrated Catalog)

The Working Group on Information Systems and Services) is a subsidiary body supporting the Committee on Earth Observing Satellites (CEOS). WGISS promotes collaboration in the

development of systems and services that manage and supply these observatory data.

CEOS IDN (CEOS International Directory Network)

An international effort developed to assist researchers in locating information on available datasets and services. The directory is sponsored as a service to the Earth science community.

FedEO (Federated Earth Observation Missions)

FedEO provides interoperable access, following ISO/OGC interface guidelines, to Earth Observation metadata.

OGC (Open Geospatial Consortium)

OGC is an international industry consortium of private companies, government agencies and universities participating in a consensus process to develop publicly available interface standards.

1.4 Specifications and references

(1) OpenSearch specifications and extensions

OpenSearch implementations within CEOS are encouraged to comply with the following OGC standard specifications.

- OpenSearch Geo and Time extension [OGC 10-032r8]
- OpenSearch Extension for EO Products [OGC 13-026r5]
- OpenSearch Parameter Extension - Draft 2

Conformance to these specifications helps achieve consistent level of implementation across CEOS agencies and systems.

Reading these specifications may require understanding of the following related specification as well.

- O&M EOP Profile [OGC 10-157r4]

See below for the brief description of the specifications and related extensions.

[OpenSearch Specification, 1.1, Draft 5](#)

http://www.opensearch.org/Specifications/OpenSearch/1.1/Draft_5

[OpenSearch Geo and Time Extensions \[OGC 10-032r8\]](#)

The Geo and Time Extensions document specifies a series of parameters that can be used to geographically constrain search results including bounding box, geometry, start and end of a temporal extent. It also defines a set of fields to be used to express geographical and temporal context in the search results.

OpenSearch Extension for EO Products [OGC 13-026r5]

This extension defines query parameters that allow the filtering of search results with the fields that are unique to EO products, eg. platform(satellite), sensor, processing center, etc. The OpenSearch query parameters defined in this document are aligned with O&M EO Profile [OGC 10-157r4] that describes EO products metadata.

The specification can be found here:

https://portal.opengeospatial.org/files/?artifact_id=61006&version=1

OpenSearch Parameter Extension - Draft 2

The OpenSearch Parameter Extension sets rules to describe valid range and entries for query parameters in OSDD. This extension is not an OGC standard.

The specification can be found here:

http://www.opensearch.org/Specifications/OpenSearch/Extensions/Parameter/1.0/Draft_2

OpenSearch Relevance extension - Draft 1

http://www.opensearch.org/Specifications/OpenSearch/Extensions/Relevance/1.0/Draft_1

O&M (Observation and Measurement) EO Profile [OGC 10-157r4]

This standard defines the encodings required to describe metadata of Earth Observation (EO) data.

The O&M standard (ISO 19156 and OGC 10-025) defines a conceptual schema encoding for observations, and for features involved in sampling when making observations.

(2) Reference

Reference documentations are as follows.

- ATOM syndication format (IETF RFC4287)
<http://tools.ietf.org/html/rfc4287>
- FEDEO Client Partner Guide, Issue 1, Revision 0, 05/02/2014, HMASE-SPB-D3200.4.
- CWIC OpenSearch Best Practices
<https://wiki.earthdata.nasa.gov/download/attachments/28541363/CWIC%20Open%20Search%20Best%20Practices.docx?version=3&modificationDate=1394136508686&api=v2>
- CWIC OpenSearch Data Partner Guide
http://ceos.org/document_management/Working_Groups/WGISS/Projects/CWIC/OpenSearch/CWIC_OpenSearch_Data-Partner-Guide.doc

- CWIC OpenSearch Client Partner Guide
http://ceos.org/document_management/Working_Groups/WGISS/Projects/CWIC/OpenSearch/CWIC_OpenSearch_Client-Guide.doc

1.5 What is OpenSearch ?

OpenSearch is a collection of simple formats for the sharing of search results. The OpenSearch specification defines description document (OSDD) format that can be used to describe a search engine so that it can be used by search client applications.

The OpenSearch response elements can be used to extend existing syndication formats, such as RSS and Atom, with the extra metadata needed to return search results. Figure 3 shows a diagram of client and server interaction of OpenSearch.

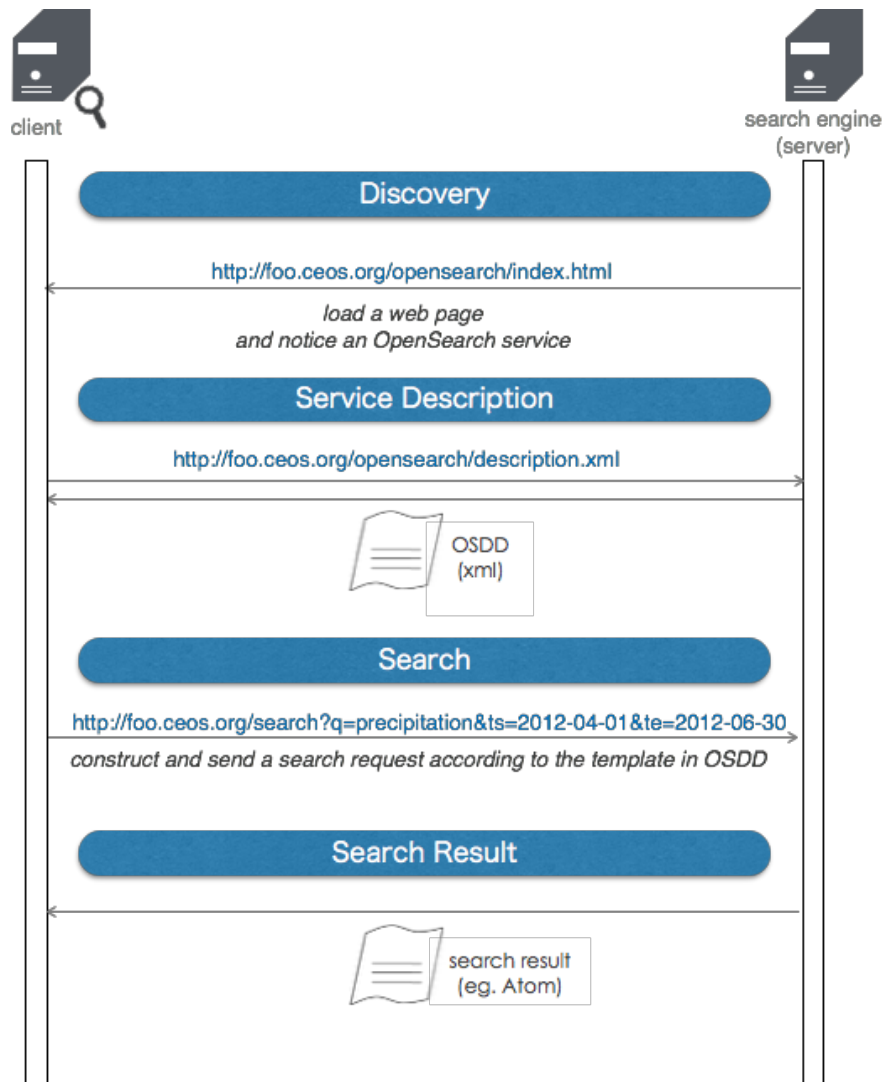


Figure 3 - OpenSearch client-server interaction

Discovery

A client can learn the support of OpenSearch service of the server by looking at the html header of its web page. A link element with a rel attribute equal to “search” shows the location of the OSDD in the value of “href” attribute as shown below.

```
<html>
<head>
<link rel="search"
type="application/opensearchdescription+xml"
title="Foo Search Engine"
href="http://foo.ceos.org/xxx/description.xml">    ... URL to OSDD
...
...
```

Description of Service

The client uses the OSDD to learn about the public interface of the server. The OSDD contains parameterized URL templates that indicate how the search client should make search requests. The URL template can iterate as many times as the output format of the search result are supported.

```
<OpenSearchDescription
xmlns="http://a9.com/-/spec/opensearch/1.1/"
xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
xmlns:atom="http://www.w3.org/2005/Atom"
xmlns:time="http://a9.com/-/opensearch/extensions/time/1.0/"
xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/"
xmlns:eo="http://a9.com/-/opensearch/extensions/eo/1.0/"
xmlns:parameters="http://a9.com/-/spec/opensearch/extensions/parameters/1.0/"
xmlns:dc="http://purl.org/dc/elements/1.1/">
...
<description>This service is ...</description>
...
<Url type="application/atom+xml" rel="results"
template="http://...search/atom?q={searchTerms}&ts={time:start}&te={time:end}"
"/>
<Query role="example" searchTerms="water"
time:start="2012-04-01"
time:end="2012-06-30"/>
...
```

```
</OpenSearchDescription>
```

Search

A client constructs the request URL by replacing each {xxx} in the URL template with a value and sends it to the server via HTTP GET. For example :

```
<Url type="text/html"  
  template="http://...search?q={searchTerms}&ts={time:start}&te={time:end}>
```

can become something like :

```
http://...search?q=water&ts=2012-04-01&te=2012-06-30
```

to be a valid search request string.

Search Response

OpenSearch doesn't define or require any specific encoding format for the search response. Instead, it defines a set of search-related metadata elements which can be inserted into existing encoding formats. Typically, list-based XML syndication formats - such as RSS 2.0 and Atom 1.0 - are used.

The OpenSearch response elements include :

totalResults	:	number of total results
startIndex	:	index number corresponding to the first entry item returned
itemsPerPage	:	number of results returned
Query	:	search query to get the same search response

Below is an example of a search response in Atom, to the free keyword “water”. Notice the three OpenSearch elements appears inside the feed element.

```
<feed xmlns=http://www.w3.org/2005/Atom
xmlns:os="http://a9.com/-/spec/opensearch/1.1/">
<title>OpenSearch response to" water"</title>
<id>http://example.ceos.org/search?q=water</id>
<author><name>Taro Yamada</name></author>
<os:totalResults>231</os:totalResults>
<os:startIndex>51</os:startIndex>
<os:itemsPerPage>50</os:itemsPerPage>
<os:Query role="request" searchTerms="water" time:start="2012-04-01"
time:end="2012-06-30" />
<entry>
  <title>Water Resource Management Guideline</title>
  <id>http://aaa.ceos.org/waterResource/guideline/v001</id>
  <link href="http://aaa.ceos.org/waterResource/guideline/v001"/>
</entry>
<entry>
  ...
</entry>
<entry>
  ...
  </entry>
<entry>
  ...
  </entry>
</feed>
```

2 CEOS OpenSearch Best Practices

This chapter goes through best practices that have been worked out from OpenSearch server implementations at CEOS agencies and projects.

The list of best practices are summarised in the table below.

ID	Category	Title	Classification
CEOS-BP-001	Two step search	Support of two step search	Recommended
CEOS-BP-002	OSDD	Support of OpenSearch Parameter Extension (Draft 2)	Recommended
CEOS-BP-003		rel attribute of the URL in OSDD	Recommended
CEOS-BP-004	BP Compliance Declaration	CEOS OpenSearch Best Practice identifier	Recommended
CEOS-BP-005	Search Request	Supported search parameters	Requirement
CEOS-BP-006		Multi-words for searchTerms	Recommended
CEOS-BP-007		Use of startPage over startIndex	Recommended
CEOS-BP-008		Search with geo:name	Recommended
CEOS-BP-009		Output encoding format in search URL	Optional
CEOS-BP-010	Search Response	Output encoding format support (atom)	Requirement
CEOS-BP-011		Support of dc:identifier in search response to allow for search-by-ID	Requirement
CEOS-BP-012		Metadata representation in search response	Recommended
CEOS-BP-013		atom:summary	Recommended
CEOS-BP-014		GeoRSS	Recommended
CEOS-BP-015		Browse Image	Recommended
CEOS-BP-016		Data access	Recommended
CEOS-BP-017	Exceptions	Exception codes	Recommended

Table 1 - List of best practices

2.1. Two Step Search

CEOS-BP-001 - Support of two step search [Recommended]

One serious hurdle to overcome in searching for data is the great number of data items to account for in responses, as well as the expected number of successful “hits” for a query. In ordinary web searches, the searcher is usually looking for a small number of web pages or documents. Relevance ranking typically does a good job of presenting these successful hits

near the top of the returned list, followed by single point-and-click retrievals. However, when searching for Earth science data covering large time periods or spatial areas, a user will often specify a set of constraints to find an appropriate data collection together with space-time criteria for files within that data collection. Often, the precision of the data collections returned for the search is low, with many spurious hits. However, the space-time precision of the files is often quite high: that is, the user truly wants to use all the data files of a desirable data collection set that fall within the space-time region of interest. Thus, searching for all data satisfying both dataset content and space-time region at the same time can produce a great many spurious hits, i.e., all the files for data collections that are *not* desired.

The 2 step search consists of a collection level search and the subsequent granule level search (or file-level search) as Figure 4 shows.

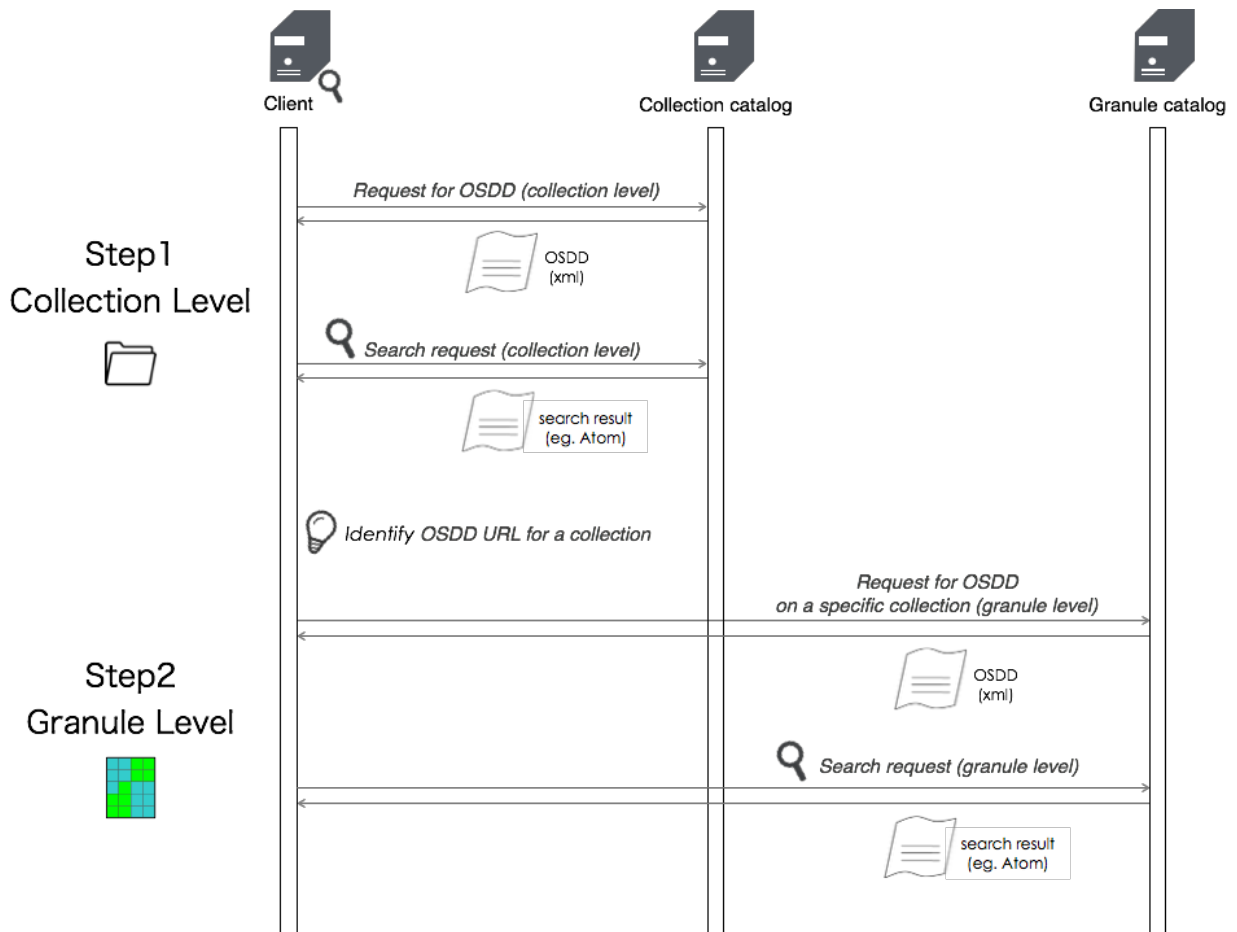


Figure 4 - 2 Step Search Diagram

In order to provide a well-defined search path from a collection of interest to granules associated with that collection, we advocate the use of two-step searching leveraging the following,

1. link elements of relation search (rel='search') within collection entries. These links point to a granule-level OSDD specific to the collection entry
2. Granule level OSDDs that can be tailored to a specific collection

The key requirement of this approach is an identifier that **both** the collection and granule resource understand as uniquely describing the collection.

The advantages of this approach are as follows,

- A client can navigate from dataset to granule with only an understanding of the OpenSearch Specification.
- A server can link between datasets and granules using any set of unique identifiers they choose. For example, identifier plus version
- It allows the client to determine what search parameters are available to the user at the granule level using the OSDD

For example, a collection level search result looks something like this :

```
<feed>
  ...
  <entry>
    <title>Collection A</title>
    <link rel="search" type="application/opensearchdescription+xml"
href="http://foo.ceos.org/search/osdd.xml?collectionId=collection\_a" />
    ...
  </entry>

  <entry>
    <title>Collection B</title>
    <link rel="search" type="application/opensearchdescription+xml"
href="http://foo.ceos.org/search/osdd.xml?collectionId=collection\_b" />
    ...
  </entry>
  ...
</feed>
```

The value of href attribute of the link element with the rel attribute equal to “search”, in this case, http://foo.ceos.org/search/osdd.xml?collectionId=collection_a, for the collection A, is the URL to the OSDD specific to collection A.

Looking for the OSDD in each result entry, a client can be navigated to URL template(s) usable for the collection’s granule level search.

The URL template for the granule search contains a search parameter dedicated to the specific collection, which is hard-coded and not meant to be replaced by the client when

constructing a query. Eg.

```
<Url type="application/atom+xml" rel="results"
template="http://../search/collection.atom?collectionId=collection_a&q={searchTerms}&ts={time:start}&te={time:end}&bbox={geo:box}>
```

Notice “collection_a” is not enclosed in braces.

2.2 OSDD

CEOS-BP-002 - Support of OpenSearch Parameter Extension (Draft 2)

[Recommended]

The Parameter Extension explicitly advertises the valid lists and ranges of search parameters. Thus, it allows to reduce ambiguity and error dramatically.

CEOS recommends to support the OpenSearch Parameter Extension with the following constraints :

- the “value” attribute should always be set to facilitate the client task (i.e. avoid to parse the <Url> templates)
- other attributes (i.e. “minimum”, “maximum”) and options elements should be provided only when it makes sense

For conveying the server’s support condition on free keywords (searchTerms) and geometry, the server should declare adherence to a specific, existing standard within the parameter definition using the Parameter Extension.

Note that the parameter extension should be provided

See below for the examples.

Free Text Keyword

```
<param:Parameter name="q" value="{searchTerms}">
  <atom:link rel="profile"
href="http://lucene.apache.org/core/2_9_4/queryparsersyntax.html" title="This
parameter follows the Lucene free text search implementations" />
</Parameter>
```

It tells a client to see the href link for the notation rule of free keywords (q) the server complies with. In this case, one can know Apache Lucene is used.

Geometry

```
<Parameter name="geometry" value="{geo:geometry}">
  <atom:link rel="profile" href="http://www.opengis.net/wkt/LINESTRING"
  title="This service accepts WKT LineStrings"/>
  <atom:link rel="profile" href="http://www.opengis.net/wkt/POINT"
  title="This service accepts WKT Point"/>
</Parameter>
```

This explains that a server supports point and line among other geometry types, and the convention can be found in the value of href.

CEOS-BP-003 - rel attribute of the URL in OSDD [Recommended]

CEOS recommends the use of rel attribute of Url element in OSDD in the following manner¹.

- rel="collection" to advertise collection level search query URL
- rel="results" (default) to mean granule level search query URL

If rel attribute is not specified, it is assumed that its value is "results".

CEOS-BP-004 - CEOS OpenSearch Best Practice identifier [Recommended]

CEOS recommends to add the string "CEOS-OS-BP-V1.0" within the <Tags> element of the OSDD. This should be used by a CEOS OpenSearch client to detect from the OSDD that a server conforms to this specification.

2.3 Search Request

CEOS-BP-005 - Supported search parameters [Requirement]

OpenSearch Geo and Time extension [OGC 10-032r8] gives a list of fundamental search parameters sufficient to constrain the search by time and space. And OpenSearch Extension for EO Products [OGC 13-026r4] complements it with optional parameters specific to EO data.

From these specifications, an OpenSearch CEOS implementation should support the

¹ As per <http://www.ietf.org/rfc/rfc6573.txt>.

following minimum set of search parameters for both “collection” and “granule” levels :

- *count*
- *searchTerms* (optional for “granule level” search)
- *startPage* or *startIndex*
- *geo:uid*
- *geo:box*
- *time:start*
- *time:end*

Note on {geo:geometry} parameter : the option of polygon in the geometry parameter raises practical issues with HTTP GET requests. Polygons can contain an arbitrary number of vertices which, when expressed as a constraint in a HTTP GET request, may exceed the URL length limits of some HTTP servers. This would suggest that HTTP POST should also be supported in some instances. However, this is not recommended since it would 1) add an unwanted degree of complexity to the specification and 2) using POST to request data would violate a RESTful implementation - POST should be used to “create” a resource on the server not to get a “resource”. As a consequence CEOS recommends the use of HTTP GET only for search request with appropriate HTTP error message (i.e. HTTP 414) when URI is too long.

CEOS-BP-006 - Multi-words for searchTerms [Recommended]

OpenSearch specifications do not define or suggest any convention for the notation of multiple keywords used in searchTerms parameter. However, consistent notation rule for expressing a phrase and logical operators are important from a user’s (i.e. client’s) perspective.

For example, what is the right way to make a search for “air temperature” as a phrase ? One can imagine a client could be bewildered with an enormous results hits when the query was interpreted as “air OR temperature” at the server, for example.

Turning to the parameter extension can resolve this by the server advertising the rule of searchTerms in OSDD. (See 2.2 for details)

And when without the parameter extension, following interpretation comes into effect as default.

- white space-delimited words not enclosed in double quote “ “ represents logical AND
(e.g.) q=air temperature : air AND temperature
- whitespace delimited words enclosed in double quote “ “ represents phrase

(e.g.) q="air temperature" : "air temperature" in one phrase
Note the examples are not URL encoded.

CEOS-BP-007 - Use of startPage over startIndex [Recommended]

OpenSearch specification allows the use of both startIndex and startPage for pagination control.

It is recommended to use startPage over startIndex to better match to web application experience.

And startPage should precede when both startPage and startIndex are supported and specified in a search by a client.

CEOS-BP-008 - Search with geo:name [Recommended]

It is recommended to support geo:name and optionally, geo:radius. If geo:name is supported but geo:radius is not, it is recommended the search be treated as point search.

The geo:name can be resolved by Gazetteer and the most common one used by CEOS agencies is geonames.org (<http://api.geonames.org/searchJSON>).

The geo:radius default/not-default value should be indicated by the server in the OSDD within the "title" attribute of the corresponding Parameter element (example below)

```
<parameters:Parameter name="radius" value="{geo:radius}" title="Expressed in meters. Default value is 10000"/>
```

CEOS-BP-009 - Output encoding format in search URL [Optional]

CEOS recommends the appearance of encoding format as a resource extension in the search URL. Note the "type" attribute of the URL element identifies the encoding format supported by a service, but this practice complements it considering shareable URL (i.e. easier to bookmark).

```
http://foo.ceos.org/search/collection.atom?collectionId=Landsat_8&startPage=1&count=10&timeStart=&timeEnd=&geoBox=
```

2.4 Search Response

CEOS-BP-010 - Output encoding format support [Requirement]

CEOS OpenSearch implementations are recommended the support of Atom as a mandatory format to be supported, as OpenSearch Geo and Time Extension [10-032r8] requires.

CEOS-BP-011 - Support of dc:identifier in search response to allow for search-by-ID [Requirement]

The dc:identifier element in the search result allows navigating from search results to a single entry inside these results using the OpenSearch parameter {geo:uid} as described in [OGC 10-032r8].

This makes possible the use cases :

- Put a link on a Web site pointing to a single (OpenSearch) catalog item (using a URL) to illustrate a particular event (e.g. an earthquake in the Himalaya)
- End users being able to bookmark and retrieve a single item

CEOS-BP-012 - Metadata representation in search response [Recommended]

Although Atom format allows any foreign (or user-defined) elements to be included and OpenSearch Extension for EO Products [OGC 13-026r5] along with O&M EOP Profile [OGC 10-157r4] suggests many options for metadata fields, CEOS OpenSearch implementations are encouraged to rely on `atom:link@rel="alternate"` or `atom:link@rel="via"` for detailed representation of the metadata. (The "via" relation should be preferred to convey the authoritative resource or the source of the information from where the atom:entry is made.) In the same context, `atom:link@rel="describedBy"` should be used to reference to the documentation (a file with human-readable information about the resource).

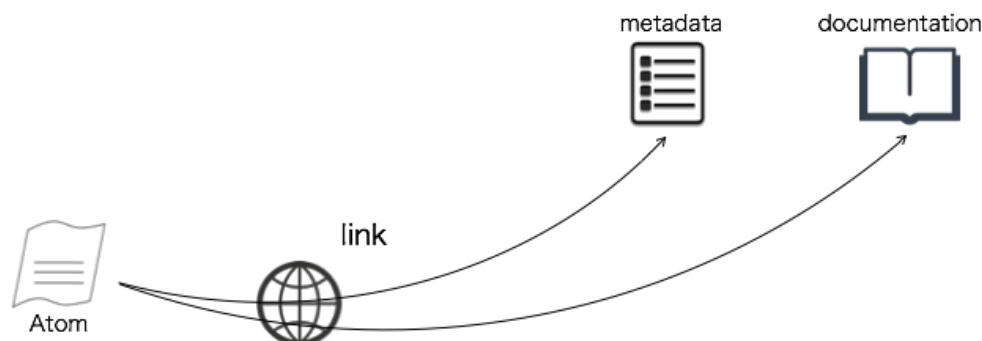


Fig5 - Referencing Metadata and Documentation

Eg.

```
<link rel="alternate" href="http://foo.ceos.org/foo/dataset_abc.xml"
type="application/vnd.iso.19139+xml"/>
<link rel="describedBy" href="http://foo.ceos.org/foo/dataset_abc.html"
type="text/html"/>
```

Table 2 defines the MIME Types to be used to reference various metadata format

Metadata format	MIME Type
ISO 19139	application/vnd.iso.19139+xml
OGC 10-157r3 OGC 10-157r4	application/gml+xml
Dublin core ²	application/xml
Markdown	text/x-markdown

Table 2 - MIME types for metadata format

CEOS-BP-013 - atom:summary [Recommended]

Default atom:summary is "text" - i.e. if MIME type is not specified then client should assumed that it is of type "text". Otherwise MIME type should be specified according to Table 2. In case of type "text/html", it needs to be correctly escaped. CEOS OpenSearch implementations are recommended to have "useful" information in the actual metadata and not only in atom:summary. Canonical source of information should be in the metadata (atom:link or feed/entry) and not in the atom:summary.

CEOS-BP-014 - GeoRSS [Recommended]

For representing geographical extent in GeoRSS in the search result, it is recommended when applicable to use "GeoRSS Simple" over "GeoRSS GML".

² According to the XML schema <http://www.loc.gov/standards/sru/recordSchemas/dc-schema.xsd>

When non applicable (e.g. footprint made of multiple polygons), “GeoRSS GML” should be used.

CEOS-BP-015 - Browse Image [Recommended]

CEOS OpenSearch implementations are recommended to provide a URL to the granule’s browse image when available, by either atom:link@rel=”icon” or Media RSS (eg. media:content/media:Category=QUICKLOOK).

CEOS-BP-016 - Data access [Recommended]

CEOS OpenSearch implementations are recommended the search result should use link element with an attribute rel=”enclosure” along with a MIME type attribute type=”xxx” for embedding data access URL.

2.5 Exceptions

CEOS-BP-017 - Exception codes [Recommended]

OpenSearch Geo and Time Extensions [OGC 10-032r8] recommends the use of HTTP status codes as following, 4xx for client errors, and 5xx for server errors.

- 400 Bad Request: The request has an invalid syntax (i.e. badly formatted geometry)
- 413 Request Entity Too Large: The request originates too many returnable hits
- 415 Unsupported media type : Media type in the request is not available or valid.
- 500 Internal Server Error: Default code for the server side for an execution error.
- 501 Not Implemented: When requesting an unimplemented feature (e.g. relation operator not supported).
- 503 Service Unavailable: When the search service is temporarily not available (due to overload or other reasons).
- 504 Gateway Timeout: When the search engine is a broker or aggregator to other services that fail to produce an answer within a giving time frame.

CEOS OpenSearch implementations are recommended to support these codes.

2.6 Future discussions

Future discussion topics include the following.

- ranking / relevance score
- aggregation of multiple source results at client side

- recommended Schema.org itemtype/itemprop to be used for granule metadata and collection metadata

3. Closer look on implementations

This chapter shows comparison among implementations : CWIC, FedEO, and CNES. (other implementations may be added in the future)

Note: CNES endpoints will not be available until january 2015

3.1 Basics

OSDD URL

Collection Level

IDN/CWIC	<a href="http://gcmd.nasa.gov/KeywordSearch/default/openSearch.jsp?Portal=cwic&clientId=<your client id here>">http://gcmd.nasa.gov/KeywordSearch/default/openSearch.jsp?Portal=cwic&clientId=<your client id here>
FedEO	http://fedeo.esa.int/opensearch/description.xml (rel="collection")
CNES	http://theia.cnes.fr/resto2/api/collections/describe.xml

Granule Level (examples)

IDN/CWIC	<ul style="list-style-type: none"> • http://cwic.wgiss.ceos.org/opensearch/datasets/osdd.xml?clientId=cwicClient • http://cwic.wgiss.ceos.org/opensearch/datasets/Landsat_8/osdd.xml?clientId=cwicClient
FedEO	<ul style="list-style-type: none"> • Option 1 (specific) http://fedeo.esa.int/opensearch/description.xml?parentIdentifier={dataset-id} • Option 2 (general) http://fedeo.esa.int/opensearch/description.xml (rel="results" - default) • Example 1 http://fedeo.esa.int/opensearch/description.xml?parentIdentifier=EOP%3AVITO%3AVGT_S10 • Example 2: http://fedeo.esa.int/opensearch/description.xml?parentIdentifier=MIR_SC_F0
CNES	http://theia.cnes.fr/resto2/api/collections/{collection}/describe.xml where {collection} should be replaced by the name of the collection (e.g. Landsat)

Search URL

Collection Level

IDN/CWIC	http://gcmd.gsfc.nasa.gov/KeywordSearch/OpenSearch.do?searchTerms=Landsat*&MetadataType=0&output=atom&Portal=cwic&clientId=fromgcmd
FedEO	http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2Bxml&query=landsat Note: parentIdentifier=EOP:ESA:FEDEO is default value for EOP:ESA:FEDEO "parentIdentifier" and can be added to request above. application/atom+xml is the default value for the httpAccept parameter.
CNES	http://theia.cnes.fr/resto2/api/collections/search.atom?q=Landsat

Granule Level (examples)

IDN/CWIC	http://cwic.wgiss.ceos.org/opensearch/granules.atom?datasetId=Landsat_8&startPage=1&count=10&timeStart=&timeEnd=&geoBox=&clientId=cwicClient
FedEO	http://fedeo.esa.int/opensearch/request/?httpAccept=application%2Fatom%2Bxml&parentIdentifier=EOP%3ASPOT%3AMULTISPECTRAL_10m&startRecord=1&startDate=2013-09-13T00%3A00%3A00Z&endDate=2013-09-16T10%3A00%3A59Z&instrumentShortName=HRGNb1&bbox=3.711001014709498%2C44.549828243256%2C22.617251014709005%2C52.987328243256
CNES	http://theia.cnes.fr/resto2/api/collections/Landsat/search.atom?lang=en&q=images%20acquired%20in%20may%202013%20over%20Paris

Documents

IDN/CWIC	<ul style="list-style-type: none"> • CWIC Client Partner Guide (OpenSearch) v0.9 • CWIC Data Partner's Guide (OpenSearch) v0.9 • CWIC Software Exception Handling v1.2
FedEO	<ul style="list-style-type: none"> • FEDEO Client Partner Guide (v1.0)
CNES	<ul style="list-style-type: none"> • http://github.com/jjrom/resto2

3.2 Remarkable Practices

OSDD

Support of parameter extension

IDN/CWIC	YES Note: Recommended attributes are “name”, “value”, “title”. “uiDisplay” will be an optional attribute in the future.
FedEO	YES
CNES	YES

Support of Explain operation (SRU)

The Explain document (from the OASIS SearchRetrieve standard) provides additional metadata about the search endpoint, the parameters accepted by the search endpoint and the possible response schemas. It fulfills a similar role as an OGC Capabilities document.

IDN/CWIC	NO
FedEO	YES
CNES	NO

Output encoding format in a search URL

IDN/CWIC	resource extension (eg. description.atom) and content negotiation
FedEO	a search parameter (httpAccept). This search parameter is borrowed from the OASIS SearchRetrieve (and SRU) standard. Note : Other ways i.e. content-negotiation using HTTP header (httpAccept) and resource extension will be supported in the future. Q2 2015)
CNES	resource extension (e.g. search.atom) and content negotiation

Support of 2 step search

IDN/CWIC	YES
FedEO	YES
CNES	YES

Search Parameters

Supported Search Parameters

C : Supported in Collection Level search parameter,

G : Supported in Granule Level search parameter

Superscript : M : Mandatory, O : Optional, 1 : dependent on subsidiary catalog

Namespaces prefixes : Table 2 of OGC 10-032r8 and Table2 of OGC 13-026r5.

Search Parameters	IDN/CWIC	FedEO	CNES
count	CG	C	G
searchTerms	C	CG ¹	CG
startIndex		C	CG
startPage	O	CG	CG
dc:publisher		C	
dc:subject		CG ¹	
dc:title		C ¹	
dc:type		C ¹	
time:start	CG	CG	G
time:end	CG	CG	G
geo:box	CG	CG	G
geo:geometry			
geo:lat		CG	G
geo:lon		CG	G
geo:name		CG	G
geo:radius		CG	G
geo:uid		CG	G
eo:acquisitionStation		C ¹ G ¹	
eo:acquisitionType		G ¹	
eo:acquisitionSubType		G ¹	
eo:antennaLookDirection		G ¹	
eo:archivingCenter		G ¹	
eo:cloudCover		G ¹	G

Search Parameters	CWIC	FedEO	CNES
eo:completionTimeFromAscendingNode		G ¹	
eo:compositeType		G ¹	
eo:dopplerFrequency		G ¹	
eo:frame		G ¹	
eo:illuminationAzimuthAngle		G ¹	
eo:illuminationElevationAngle		G ¹	
eo:imageQualityDegradation		G ¹	
eo:incidenceAngleVariation		G ¹	
eo:instrument		CG ¹	G
eo:maximumIncidenceAngle		G ¹	
eo:minimumIncidenceAngle		G ¹	
eo:orbitDirection		G ¹	
eo:orbitNumber		C ¹ G ¹	G
eo:orbitType		G ¹	
eo:organisationName		C	G
eo:parentIdentifier	G (datasetId)	CG ¹	
eo:platformSerialIdentifier		G ¹	
eo:platform		CG ¹	G
eo:polarisationChannels		G ¹	
eo:polarisationMode		G ¹	
eo:processingCenter		C ¹ G ¹	
eo:processingLevel		G ¹	G
eo:processorName		G ¹	
eo:productionStatus		G ¹	
eo:productType		CG ¹	G
eo:resolution		G ¹	G
eo:sensorType		G ¹	
eo:sensorMode		G ¹	
eo:snowCover		G ¹	G
eo:startTimeFromAscendingNode		G ¹	
eo:swathIdentifier		G ¹	

Search Parameters	CWIC	fedEO	CNES
eo:track		C ¹	
semantic:classifiedAs		C	
sru:recordSchema		CG	

Supported output format

IDN/CWIC	ATOM, RSS ¹ , HTML ¹ [1] IDN only
FedEO	ATOM, RSS ¹ , SRU ¹ , RDF ¹ , HTML ² [1] depends on data partner [2] planned for Q2 2015
CNES	ATOM, GeoJSON, HTML

Output schema support

IDN/CWIC	NO
FedEO	YES (dc, iso, O&M EOP v1.0 and v1.1)
CNES	NO

Free Keyword notation (searchTerms)

IDN/CWIC	Conveyed by Parameter Extension (Parameter/link@rel="profile")
FedEO	whitespace for AND double quote for phrase
CNES	Natural language queries (e.g. "images over Paris between may and june 2013")

Pagination

IDN/CWIC	startPage, count
FedEO	startIndex, startPage, count

CNES	startIndex, startPage, count
------	------------------------------

Search Result

Browse

IDN/CWIC	link@rel="icon"
FedEO	link@rel="icon" also supports Media RSS notation (use of media:content/media:Category=QUICKLOOK)
CNES	link@rel="icon" also supports Media RSS notation (use of media:content/media:Category=QUICKLOOK)

Thumbnail

IDN/CWIC	N/A
FedEO	Media RSS notation (use of media:content/media:Category=THUMBNAIL)
CNES	Media RSS notation (use of media:content/media:Category=THUMBNAIL)

Mask Image

IDN/CWIC	N/A
FedEO	Media RSS notation (use of media:content/media:Category=MASK)
CNES	N/A

Metadata

IDN/CWIC	link@rel="via"
FedEO	link@rel="alternate" or link@rel="via"
CNES	link@rel="via"

Documentation

IDN/CWIC	link@rel="describedBy"
FedEO	N/A
CNES	N/A

Data

IDN/CWIC	link@rel="enclosure"
FedEO	link@rel="enclosure"
CNES	link@rel="enclosure"

Error Handling

Exception Status Code (4xx, 5xx)

IDN/CWIC	Compliant with OGC OpenSearch [OGC 10-032r8]
FedEO	Compliant with OGC OpenSearch [OGC 10-032r8] and OGC OWS Common [OGC 06-121]
CNES	Compliant with OGC OpenSearch [OGC 10-032r8]

APPENDIX A - List of OpenSearch Endpoints

This appendix shows OpenSearch Endpoints of CEOS agency catalogs.

FedEO : <http://fedeo.esa.int/opensearch/description.xml>
IDN/CWIC : <http://gcmd.gsfc.nasa.gov/KeywordSearch/default/openSearch.jsp?Portal=cwic&clientId=<your client id here>>
NASA ECHO : https://api.echo.nasa.gov:443/opensearch/datasets/descriptor_document.xml?clientId=xxxxxxx
CNES : <http://theia.cnes.fr/api/collections/describe.xml>

APPENDIX B - Examples

This appendix contains example code snippets from CEOS agencies' implementations.

B1. OSDD (Collection Level)

IDN/CWIC

```
http://gcmd.nasa.gov/KeywordSearch/default/openSearch.jsp?Portal=cwic&clientI  
d=<your client id here>
```

FedEO

```
http://fedeo.esa.int/opensearch/description.xml
```

CNES

```
http://theia.cnes.fr/resto2/api/collections/describe.xml
```

B2. OSDD (Granule Level)

IDN/CWIC

```
http://cwic.wgiss.ceos.org/opensearch/datasets/Landsat_8/osdd.xml?clientId=cwicClient
```

FedEO

```
http://fedeo.esa.int/opensearch/description.xml?parentIdentifier=EOP:VITO:VGT_S10
```

CNES

```
http://theia.cnes.fr/resto2/api/collections/Landsat/describe.xml
```

B3. Search request

IDN/CWIC

```
http://cwic.wgiss.ceos.org/openserach/granules.atom?datasetId=INPE_CBERS2B_CC  
D&startIndex=2&count=1&timeStart=2007-09-25T00:00:00Z&timeEnd=2010-03-10T00:0  
0:00Z&geoBox=-180,-90,180,90
```

FedEO

```
http://fedeo.esa.int/openserach/request/?httpAccept=application/atom%2Bxml&pa  
rentIdentifier=EOP%3AEGEOS%3AEGEOS%23IKONOS&startDate=2000-10-31T00:00:00Z&ma  
ximumRecords=2&cloudCover=10]&endDate=2013-10-31T00:00:00Z&name=frascati&radi  
us=1000
```

CNES

```
http://theia.cnes.fr/resto2/api/collections/Landsat/search.atom?q=marseille
```

B4. Search response (Atom)

IDN/CWIC

```
<?xml version="1.0" encoding="UTF-8"?>
<feed xmlns="http://www.w3.org/2005/Atom"
xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/"
xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/"
xmlns:georss="http://www.georss.org/georss/1.0"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:time="http://a9.com/-/opensearch/extensions/time/1.0/"
xmlns:cwic="http://cwic.wgiss.ceos.org/opensearch/extensions/1.0/"
xmlns:esipdiscover="http://commons.esipfed.org/ns/discovery/1.2/">
<title>CWIC OpenSearch Response</title>
<updated>2014-09-24T08:47:15Z</updated>
<author>
  <name>CEOS WGISS Integrated Catalog (CWIC) - CWIC Contact - Email:
cwic-help@wgiss.ceos.org - Web: http://wgiss.ceos.org/cwic</name>
  <email>cwic-help@wgiss.ceos.org</email>
</author>
<id>http://cwic.wgiss.ceos.org/opensearch/granules.atom</id>
<link rel="self" type="application/atom+xml"
href="http://cwic.wgiss.ceos.org/opensearch/granules.atom?datasetId=INPE_CBER
S2B_CCD&startIndex=2&count=1&timeStart=2007-09-25T00:00:00Z&timeEnd=2010-03-10T00:00:00Z&geoBox=-180,-90,180,90&startPage=1"
title="This Request"/>
<link rel="search" type="application/opensearchdescription+xml"
href="http://cwic.wgiss.ceos.org/opensearch/granules.atom" title="Search this
resource"/>
<link rel="next" type="application/atom+xml"
href="http://cwic.wgiss.ceos.org/opensearch/granules.atom?datasetId=INPE_CBER
S2B_CCD&count=1&startPage=2&geoBox=-180,-90,180,90&timeStart=
2007-09-25T00:00:00Z&timeEnd=2010-03-10T00:00:00Z" title="Next Page"/>
<link rel="first" type="application/atom+xml"
href="http://cwic.wgiss.ceos.org/opensearch/granules.atom?datasetId=INPE_CBER
S2B_CCD&count=1&startPage=1&geoBox=-180,-90,180,90&timeStart=
2007-09-25T00:00:00Z&timeEnd=2010-03-10T00:00:00Z" title="First Page"/>
<link rel="last" type="application/atom+xml"
href="http://cwic.wgiss.ceos.org/opensearch/granules.atom?datasetId=INPE_CBER
S2B_CCD&count=1&startPage=62304&geoBox=-180,-90,180,90&timeSt
art=2007-09-25T00:00:00Z&timeEnd=2010-03-10T00:00:00Z" title="Last Page"/>
<opensearch:totalResults>62304</opensearch:totalResults>
<opensearch:startPage>1</opensearch:startPage>
<opensearch:itemsPerPage>1</opensearch:itemsPerPage>
<opensearch:Query role="request" cwic:datasetId="INPE_CBERS2B_CCD"
startPage="1" count="1" geo:box="-180,-90,180,90"
time:start="2007-09-25T00:00:00Z" time:end="2010-03-10T00:00:00Z"/>
```

```

<!-- Remote search completed in 0 seconds. -->
<entry>
  <title>INPE_CBERS2B_CCD:CB2BCCD16709120100310</title>

  <id>http://cwic.wgiss.ceos.org/opensearch/granules.atom?uid=INPE_CBERS2B_CCD:
  CB2BCCD16709120100310</id>
  <updated>2010-03-10 14:13:48</updated>
  <author>
    <name>ATUS USERS ASSISTANCE SERVICE - TECHNICAL CONTACT - Rod. Pres.
    Dutra, km 40; Cachoeira Paulista, SP 12630-000; Brazil - Email:
    atus@dgi.inpe.br - Phone: +55 (12) 3186-9226 - FAX: +55 (12) 3101 1507</name>
    <email>atus@dgi.inpe.br</email>
  </author>
  <georss:box>-52.85260 7.49748 -51.62200 8.62678</georss:box>
  <georss:polygon>7.62109 -52.85260 8.62678 -52.63070 8.50276 -51.62200
  7.49748 -51.84660 7.62109 -52.85260</georss:polygon>
  <dc:date>2010-03-10T14:09:45Z/2010-03-10T14:10:01Z</dc:date>
  <link title="Original source metadata" rel="via" type="text/xml"
  href="http://www.dgi.inpe.br/cwic/sceneid.php?sceneid=CB2BCCD16709120100310"/
  >
  <link title="Alternate metadata URL" rel="alternate" type="text/xml"
  href="http://www.dgi.inpe.br/cwic/sceneid.php?sceneid=CB2BCCD16709120100310"/
  >
  <link title="Browse image URL" rel="icon" type="image/jpeg"
  href="http://www.dgi.inpe.br/CDSR/display.php?TABELA=Browse&PREFIXO=Cbers
  &INDICE=CB2BCCD16709120100310"/>
  <link title="Granule ordering URL" rel="alternate" type="text/html"
  href="http://www.dgi.inpe.br/CDSR/cart-cwic.php?SCENEID=CB2BCCD16709120100310
  "/>
  <summary type="text">Granule metadata for
  INPE_CBERS2B_CCD:CB2BCCD16709120100310</summary>
</entry>
<!-- Atom response generation completed in 1ms. -->
</feed>
<!-- OpenSearch response completed in 847ms. -->

```

FedEO

```

<?xml version="1.0" encoding="UTF-8"?>
<feed xmlns="http://www.w3.org/2005/Atom"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:eo="http://a9.com/-/opensearch/extensions/eo/1.0/"
  xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/"
  xmlns:georss="http://www.georss.org/georss"
  xmlns:media="http://search.yahoo.com/mrss/"

```

```
xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
xmlns:sru="http://a9.com/-/opensearch/extensions/sru/2.0/"
xmlns:time="http://a9.com/-/opensearch/extensions/time/1.0/"
xmlns:wrs="http://www.opengis.net/cat/wrs/1.0">
<os:totalResults>8</os:totalResults>
<os:startIndex>1</os:startIndex>
<os:itemsPerPage>2</os:itemsPerPage>
<os:Query count="2"
eo:parentIdentifier="urn:ogc:def:EOP:EGEOS:EGEOS#IKONOS"
geo:box="12.650872,41.812231,12.675008,41.830249" geo:name="frascati"
geo:radius="1000" role="request" startIndex="1"
time:end="2013-10-31T00:00:00Z" time:start="2000-10-31T00:00:00Z"/>
<author>
<name>FEDEO Clearinghouse</name>
</author>
<generator>FEDEO Clearinghouse</generator>
<id>http://fedeo.esa.int/opensearch/request</id>
<title>FEDEO Clearinghouse - Search Response</title>
<updated>2014-09-26T20:46:32Z</updated>
<link
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2B
xml&parentIdentifier=EOP%3AEGEOS%3AEGEOS%23IKONOS&startDate=2000-10-3
1T00:00:00Z&maximumRecords=2&cloudCover=10]&endDate=2013-10-31T00
:00:00Z&name=frascati&radius=1000" rel="self"
type="application/atom+xml"/>
<link
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2B
xml&parentIdentifier=EOP%3AEGEOS%3AEGEOS%23IKONOS&startDate=2000-10-3
1T00:00:00Z&maximumRecords=2&cloudCover=10]&endDate=2013-10-31T00
:00:00Z&name=frascati&radius=1000&startRecord=1" rel="first"
type="application/atom+xml"/>
<link
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2B
xml&parentIdentifier=EOP%3AEGEOS%3AEGEOS%23IKONOS&startDate=2000-10-3
1T00:00:00Z&maximumRecords=2&cloudCover=10]&endDate=2013-10-31T00
:00:00Z&name=frascati&radius=1000&startRecord=3" rel="next"
type="application/atom+xml"/>
<link
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2B
xml&parentIdentifier=EOP%3AEGEOS%3AEGEOS%23IKONOS&startDate=2000-10-3
1T00:00:00Z&maximumRecords=2&cloudCover=10]&endDate=2013-10-31T00
:00:00Z&name=frascati&radius=1000&startRecord=7" rel="last"
type="application/atom+xml"/>
<link href="http://fedeo.esa.int/opensearch/description.xml" rel="search"
type="application/opensearchdescription+xml"/>
<entry>
```



```

<id>http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2Bxml&parentIdentifier=urn%3Aogc%3Adef%3AEOP%3AEGEOS%3AEGEOS%23IKONOS&uid=urn%3Aogc%3Adef%3AEOP%3AEGEOS%3AEGEOS%23IKONOS%3A20010814095358600000116244062000010418104THC&recordSchema=server-choice</id>
<link
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2Bxml&parentIdentifier=urn%3Aogc%3Adef%3AEOP%3AEGEOS%3AEGEOS%23IKONOS&uid=urn%3Aogc%3Adef%3AEOP%3AEGEOS%3AEGEOS%23IKONOS%3A20010814095358600000116244062000010418104THC&recordSchema=server-choice" rel="alternate"
title="URL for retrieving full details of the product:
urn:ogc:def:EOP:EGEOS:EGEOS#IKONOS:20010814095358600000116244062000010418104THC" type="application/atom+xml"/>
<published>2001-08-14T09:53:58Z</published>
<dc:identifier>urn:ogc:def:EOP:EGEOS:EGEOS#IKONOS:20010814095358600000116244062000010418104THC</dc:identifier>
<title>urn:ogc:def:EOP:EGEOS:EGEOS#IKONOS:20010814095358600000116244062000010418104THC</title>
<updated>2014-09-26T20:46:32Z</updated>
<summary type="html"><![CDATA[
<table>
<tr>
<td valign="top">
<a
href="http://geofuse.geoeye.com/static/browse/ikonos/2/kpms/2001/08/2001081409535860000011624406_4.jpg" target="_blank" title="View browse image">

</a>
</td>
<td valign="top">
<table>
<tr valign="top">
<td>
<b>Platform Short Name</b>
</td>
<td>IKONOS</td>
</tr>
<tr valign="top">
<td>
<b>Platform Serial Identifier</b>
</td>
<td>IKONOS-2</td>
</tr>
<tr valign="top">
<td>

```

```

<b>Sensor Type</b>
</td>
<td>OPTICAL</td>
</tr>
<tr valign="top">
<td>
<b>Sensor Operational Mode</b>
</td>
<td>PAN/MSI</td>
</tr>
<tr valign="top">
<td>
<b>Sensor Resolution</b>
</td>
<td>0.98</td>
</tr>
<tr>
<td>
<b>Media Type</b>
</td>
<td>
<a
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/atom%2B
xml&parentIdentifier=urn%3Aogc%3Adef%3Aeop%3Aegeos%3Aegeos%23IKONOS&u
id=urn%3Aogc%3Adef%3Aeop%3Aegeos%3Aegeos%23IKONOS%3A2001081409535860000011624
4062000010418104THC&recordSchema=server-choice" title="Atom
format">ATOM</a>
| <a
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/rss%2B
xml&parentIdentifier=urn%3Aogc%3Adef%3Aeop%3Aegeos%3Aegeos%23IKONOS&u
id=urn%3Aogc%3Adef%3Aeop%3Aegeos%3Aegeos%23IKONOS%3A20010814095358600000116244
062000010418104THC&recordSchema=server-choice" title="RSS format">RSS</a>
| <a
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/sru%2B
xml&parentIdentifier=urn%3Aogc%3Adef%3Aeop%3Aegeos%3Aegeos%23IKONOS&u
id=urn%3Aogc%3Adef%3Aeop%3Aegeos%3Aegeos%23IKONOS%3A20010814095358600000116244
062000010418104THC&recordSchema=server-choice" title="SRU format">SRU</a>
</td>
</tr>
<tr>
<td>
<b>Metadata</b>
</td>
<td>
<a
href="http://fedeo.esa.int/opensearch/request/?httpAccept=application/gml%2Bx

```

```

ml&parentIdentifier=urn%3Aogc%3Adef%3AEO%3AEGEOS%3AEGEOS%23IKONOS&uid=urn%3Aogc%3Adef%3AEO%3AEGEOS%3AEGEOS%23IKONOS%3A20010814095358600000116244062000010418104THC&recordSchema=server-choice" title="O&M format">O&M</a>
</td>
</tr>
</table>
</td>
<td valign="top">
<table>
<tr valign="top">
<td>
<b>Start Date</b>
</td>
<td>2001-08-14T09:53:58.00Z</td>
</tr>
<tr valign="top">
<td>
<b>End Date</b>
</td>
<td>2001-08-14T09:54:02.00Z</td>
</tr>
</table>
</td>
</tr>
</table>]]</summary>
<link
href="http://fedeo.esa.int/opensearch/request?httpAccept=application/gml%2Bxml&parentIdentifier=urn%3Aogc%3Adef%3AEO%3AEGEOS%3AEGEOS%23IKONOS&uid=urn%3Aogc%3Adef%3AEO%3AEGEOS%3AEGEOS%23IKONOS%3A20010814095358600000116244062000010418104THC&recordSchema=om" rel="alternate"
type="application/gml+xml"/>
<link
href="http://fedeo.esa.int/opensearch/request?httpAccept=application/gml%2Bxml&parentIdentifier=urn%3Aogc%3Adef%3AEO%3AEGEOS%3AEGEOS%23IKONOS&uid=urn%3Aogc%3Adef%3AEO%3AEGEOS%3AEGEOS%23IKONOS%3A20010814095358600000116244062000010418104THC&recordSchema=om11" hreflang="en" rel="alternate"
type="application/gml+xml"/>
<link
href="http://geofuse.geoeye.com/static/browse/ikonos/2/kpms/2001/08/2001081409535860000011624406_4.jpg" rel="icon" type="image/jpeg"/>
<georss:polygon xmlns:gml="http://www.opengis.net/gml">41.7496192380002
12.6632663830001 41.7518167270001 12.6084605380001 41.7540093100001
12.5561940820002 41.7563784120001 12.5010395270002 41.8224882260001
12.5004800220001 41.8887371270001 12.4994988630002 41.8864556400001
12.5547515790001 41.8844161960001 12.6068495140001 41.8822238410002

```

```

12.6620891320001 41.8157604600001 12.6632254980001 41.7496192380002
12.6632663830001</georss:polygon>
<media:group>
<media:content medium="image" type="image/jpeg"
url="http://geofuse.geoeeye.com/static/browse/ikonos/2/kpms/2001/08/2001081409
535860000011624406_4.jpg">
<media:category
scheme="http://www.opengis.net/spec/EOMPOM/1.0">QUICKLOOK</media:category>
</media:content>
</media:group>
</entry>

...
</feed>

```

CNES

```

<?xml version="1.0" encoding="UTF-8"?>
<feed xml:lang="en" xmlns="http://www.w3.org/2005/Atom"
xmlns:time="http://a9.com/-/opensearch/extensions/time/1.0/"
xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:georss="http://www.georss.org/georss"
xmlns:gml="http://www.opengis.net/gml"
xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/"
xmlns:eo="http://a9.com/-/opensearch/extensions/eo/1.0/"
xmlns:metalink="urn:ietf:params:xml:ns:metalink"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:media="http://search.yahoo.com/mrss/">
  <title></title>
  <subtitle type="html">&nbsp;|&nbsp;_pagination</subtitle>
  <generator uri="http://mapshup.com" version="2.0">resto</generator>
  <updated>2014-12-15T12:29:34+0100</updated>
  <id>7071ea35-308e-5149-a56d-ad6d8e9368d0</id>
  <link rel="self" title="self" type="application/atom+xml"
href="http://theia.cnes.fr/resto2/api/collections/Landsat/search.atom?q=marse
ille&lang=fr&maxRecords=1"/>
  <link rel="search" title="_osddLink"
type="application/opensearchdescription+xml"
href="http://theia.cnes.fr/resto2/api/collections/Landsat/describe.xml"/>
  <link rel="next" title="next" type="application/atom+xml"
href="http://theia.cnes.fr/resto2/api/collections/Landsat/search.atom?q=marse
ille&lang=fr&maxRecords=1&page=2"/>
  <os:startIndex>1</os:startIndex>

```

```

<os:itemsPerPage>1</os:itemsPerPage>
<os:Query role="request" searchTerms="marseille" language="fr" count="1"/>
<entry>
  <id>4cbcb52c-395c-5d86-862b-41ff4349b2b6</id>
  <dc:identifiant>4cbcb52c-395c-5d86-862b-41ff4349b2b6</dc:identifiant>
  <title></title>
  <published>2014-11-25T17:08:48Z</published>
  <updated>2014-11-25T17:08:48Z</updated>
  <dc:date>2014-10-30T10:17:51Z/2014-10-30T10:17:51Z</dc:date>
  <gml:validTime>
    <gml:TimePeriod>
      <gml:beginPosition>2014-10-30T10:17:51Z</gml:beginPosition>
      <gml:endPosition>2014-10-30T10:17:51Z</gml:endPosition>
    </gml:TimePeriod>
  </gml:validTime>
  <georss:where>
    <gml:Polygon>
      <gml:exterior>
        <gml:LinearRing>
          <gml:posList srsDimensions="2">43.9830355963 4.24846182664
43.9587914693 5.61895825622 42.9696032211 5.57428279738 42.9934166847
4.22715319234 43.9830355963 4.24846182664</gml:posList>
        </gml:LinearRing>
      </gml:exterior>
    </gml:Polygon>
  </georss:where>
  <link rel="alternate" type="alternate" title="Lien HTML pour
4cbcb52c-395c-5d86-862b-41ff4349b2b6"
href="http://theia.cnes.fr/resto2/collections/Landsat/4cbcb52c-395c-5d86-862b
-41ff4349b2b6.html?lang=fr"/>
  <link rel="alternate" type="alternate" title="Lien GeoJSON pour
4cbcb52c-395c-5d86-862b-41ff4349b2b6"
href="http://theia.cnes.fr/resto2/collections/Landsat/4cbcb52c-395c-5d86-862b
-41ff4349b2b6.json?lang=fr"/>
  <link rel="alternate" type="alternate" title="Lien ATOM pour
4cbcb52c-395c-5d86-862b-41ff4349b2b6"
href="http://theia.cnes.fr/resto2/collections/Landsat/4cbcb52c-395c-5d86-862b
-41ff4349b2b6.atom?lang=fr"/>
  <link rel="via" type="via" title="_metadataLink"
href="http://spirit.cnes.fr/landsat/md/LANDSAT8_OLITIRS_XS_20141030_N2A_Franc
e-MetropoleD0008H0002.xml"/>
  <link rel="enclosure" type="application/x-gzip" length="0" title="File for
4cbcb52c-395c-5d86-862b-41ff4349b2b6 product" metalink:priority="50"
href="http://theia.cnes.fr/resto2/collections/Landsat/4cbcb52c-395c-5d86-862b
-41ff4349b2b6/download"/>
  <link rel="icon" title="Browse image URL for

```

```
4cbcb52c-395c-5d86-862b-41ff4349b2b6 product"
href="http://spirit.cnes.fr/landsat/ql/LANDSAT8_OLITIRS_XS_20141030_N2A_France-MetropoleD0008H0002.png"/>
  <media:group>
    <media:content
url="http://spirit.cnes.fr/landsat/ql/LANDSAT8_OLITIRS_XS_20141030_N2A_France-MetropoleD0008H0002_thumb.png" medium="image">
      <media:category
scheme="http://www.opengis.net/spec/EOMPOM/1.0">THUMBNAIL</media:category>
    </media:content>
    <media:content
url="http://spirit.cnes.fr/landsat/ql/LANDSAT8_OLITIRS_XS_20141030_N2A_France-MetropoleD0008H0002.png" medium="image">
      <media:category
scheme="http://www.opengis.net/spec/EOMPOM/1.0">QUICKLOOK</media:category>
    </media:content>
  </media:group>
  <summary type="text">LANDSAT8/OLITIRS acquise en
2014-10-30T10:17:51Z</summary>
</entry>
</feed>
```

APPENDIX C - Search-Engine-Friendly catalog

This appendix gives some hint to make a catalog Search-Engine-Friendly with schema.org

Only ATOM format is required as an output format response. However, for implementation that support also html as an output format, it is recommended to use schema.org within the provided html response.

Additionally there are numerous ways in which that html can be designed to make it visible to search engines and, therefore, a searcher interested in earth science data.

Schema.org provides a means to expose datasets in search engine results that is explicitly geared towards earth science datasets (<https://schema.org/Dataset>). Implementors can mark up their html in a manner that allows crawlers to index their data in terms of a variety of parameters including spatial and temporal extent. For example :

```
<li itemscope itemtype='http://schema.org/Dataset'>
  <span itemprop='alternateName'>doi:10.3334/ORNLDAAC/1</span>
  <span itemprop='version' >1</span>
  <span itemprop='provider' >ORNL_DAAC</span>
  <div itemprop='description' >ABSTRACT: USGS 15 minute stream flow data for
  Kings Creek on the Konza Prairie</div>
  <span itemprop='temporal' >2000-01-01T00:00:00Z</span>
  <span itemprop='spatial' >-55.0 -180.0 90.0 180.0</span>
</li>
```

And another example :

```
<ul class='results' itemscope itemtype='http://schema.org/DataCatalog'>
  <li class="dataset" itemscope itemtype='http://schema.org/Dataset'>
    <h2 itemprop='name'>15 Minute Stream Flow Data: USGS (FIFE)</h2>
    <span class="uid">C179003030-ORNL_DAAC</span>
    <div class='dataset-identification'><span>Dataset Information:</span>
      <div class='tuple' >Short name: <span class='value short_name'
        itemprop='alternateName' >doi:10.3334/ORNLDAAC/1</span></div>
      <div class='tuple' >Version ID: <span class='value version_id'
        itemprop='version' >1</span></div>
      <div class='tuple' >Data center: <span class='value data_center'
        itemprop='provider' >ORNL_DAAC</span></div>
    </div>
    <span><div class='temporal'><span>Temporal extent:</span><div
      class='tuple'>Start: <span class='start
        value'>1984-12-25T00:00:00.000Z</span></div><div class='tuple'>Stop: <span
        class='stop value'>1988-03-04T00:00:00.000Z</span></div></div></span>
```

```

    <span class='temporal schema_org' itemprop='temporal'
>1984-12-25T00:00:00.000Z/1988-03-04T00:00:00.000Z</span>
    <span><div class='spatial'><span>Spatial extent:</span><div
class='tuple'>Bounding box: <span class='spatial-value value'>39.1 -96.6 39.1
-96.6</span></div><div class='tuple'>Point: <span class='spatial-value
value'>39.1 -96.6</span></div></span></div></span>
    <span class='spatial schema_org' itemprop='spatial' >39.1 -96.6 39.1
-96.6</span>
    <div class='dataset-identification'><span>Description:</span>
    <div class='tuple description' itemprop='description'
>ABSTRACT: USGS 15 minute stream flow data for Kings Creek on the Konza
Prairie</div>
    </div>
    <div>
    <span class='link-title'>Links:</span>
    <ul class='links'>
        <li><a class='dataset_link'
href="http://daac.ornl.gov/cgi-bin/dsvviewer.pl?ds_id=1" itemprop='url' >Data
link</a></li>
        <li><a class='dataset_link'
href="http://daac.ornl.gov/FIFE/guides/15_min_strm_flow.html" itemprop='url'
>USGS 15 minute stream flow data for Kings Creek on the Konza Prairie (VIEW
RELATED INFORMATION)</a></li>
        <li><a class='dataset_link'
href="http://gcmd.nasa.gov/getdif.htm?FIFE_STRM_15M" itemprop='url'
>doi:10.3334/ORNLDAAC/1</a></li>
        <li><a class='dataset_link'
href="https://api.echo.nasa.gov:443/catalog-rest/echo_catalog/datasets/C17900
3030-ORNL_DAAC.xml" itemprop='url' >Product metadata</a></li>
    </ul>
    </div>
    <div class="granule-search">
    <a
href="/opensearch/granules?clientId=our_html_ui&dataCenter=ORNL_DAAC&
dataset_cursor=1&dataset_id=15+Minute+Stream+Flow+Data%3A+USGS+%28FIFE%29
&dataset_number_of_results=10&shortName=doi%3A10.3334%2FORNLDAAC%2F1&
&spatial_type=bbox
&versionId=1"><span class="ui-state-default ui-corner-all
open-search-button">Granule Search</span></a>
    </div>
</li>

```