CEOS Working Group
on Information Systems and Services

# WGISS Connected Data Assets Data Partner Guide

Publication Date: 2019-04-04

Updated: 2021-12-01

Document version: V1.1

Category: WGISS Technical Document

Editors: WGISS CDA System-Level Team

# Executive Summary

**Recommendations for WCDA Data Partners**

In order to fully participate withing the CEOS/Working Group on Information Systems and Services (WGISS) Connected Data Assets (WCDA) framework as a Data Partner, there are recommendations that will make the job easier.

1. Register each distinct, searchable data set in the IDN.
2. Provide a search interface accessible via a simple URL (*i.e.,* HTTP/HTTPS GET), including parameters for starting record number and number of records desired in the response. The widely implemented OpenSearch (CEOS compliant) protocol is now mandatory.
3. Support searching on spatial bounding box.
4. Support searching on temporal extent, at least observation start and end dates.
5. Identify and, ideally, filter for limitations on search extent (spatial and temporal) to prevent search timeouts.
6. Provide search responses in well-formed text returning matching data granules.
7. Identify each returned data granule by an identifier that is unique within the inventory system.
8. Provide a capability for using the granule identifier to retrieve metadata about the granule.
9. Return URLs for browse data and direct access to granule-level data (or to a data ordering system) in the search response.

In general, these are common and widely implemented capabilities in almost any granule search system and should not represent an impediment to joining WCDA as a Data Partner.

**Table of Contents**

# 1  Introduction

This document contains the comprehensive guide for prospective WCDA Data Partners, as adopted in the CEOS WGISS Connected Data Assets project. The document introduces the architecture background and necessary skills to be a data provider, the query interface, and an implementation outline. Several detailed use cases about how to retrieve the IDN (International Directory Network) dataset ID and how to interact with the WCDA Granule Gateway servers are also included in this document.

# 2  Scope

This data partner guide applies to any data provider who plans to join WCDA. This data partner guide is currently applicable to most other potential data partners.

The target audience for this document is the community of software developers who are:

a) Implementers of IDN OpenSearch

b) Implementers of CEOS Best Practices-compliant OpenSearch servers

# 3  References

The following documents provide more background and supportive information.

| Document Reference & Version | Document Title / Description |
|---|---|
| CEOS-OPENSEARCH-BP-V1.3 | CEOS OpenSearch Best Practice Document |
| OGC 10-157r4 | Earth Observation Metadata profile of Observations & Measurements, Version 1.1 |
| OGC 10-032r8 | OpenSearch Geo and Time Extensions |
| OGC 13-026r8 | OpenSearch Extension for Earth Observation |
| OGC 07-006r1 | OpenGIS Catalogue Service Implementation Specification |

# 4  Before You Begin

This chapter introduces the background, concepts and architecture of IDN, CWIC/CMR, and FedEO, within the scope of the WCDA. The related skills you will need as a data partner are also discussed.

## 4.1 WGISS Connected Data Asset Background

For scientists who conduct multi-disciplinary research, there may be a need to search multiple catalogs in order to find the data they need. Such work can be very time-consuming and tedious, especially when different catalogs may use different metadata models and catalog interface protocols. It would be desirable, therefore, for those catalogs to be integrated into a catalog federation which will present a well-known and documented metadata model and interface protocol to users and hide the complexity and diversity of the affiliated catalogs behind the interface. With such a federation, users only need to work with the federated catalog through the public interface or API to find the data they need instead of working with various catalogs individually.

The Committee on Earth Observation Satellites (CEOS) addresses coordination of the satellite Earth Observation (EO) programs of the world's government agencies, along with agencies that receive, and process data acquired remotely from space. The Working Group on Information Systems and Services (WGISS) is a subgroup of CEOS, which aims to promote collaboration in the development of systems and services that manage and supply EO data to users world-wide. NASA's contributions to the CEOS International Directory Network (IDN) provides access to more than 50,000 Earth science data set and service descriptions (stored in the NASA Common Metadata Repository [CMR]) which cover subject areas within Earth and environmental sciences. The IDN's mission is to assist researchers, policy makers, and the public in the discovery of and access to data and related services relevant to Earth science research.

To aid in the search and discovery effort, Global Change Master Directory (GCMD) controlled keywords have been developed and are regularly being refined and expanded. These keywords are also used in other applications within the broader scientific community. Users may perform searches through the IDN website and OpenSearch API using the controlled keywords, free-text searches, map/date searches, or any combination of the above; and may also search or refine a search by data center, instrument, platform, project, or temporal/spatial resolution.

The IDN also supports draft Metadata Management Tool (dMMT), a web-based metadata authoring tool that allows metadata authors to add (or modify) data set descriptions that comply with the CMR Unified Metadata Model for Collections (UMM-C). The tool also allows metadata authors to validate and submit their metadata records directly for discovery in the IDN.
In order to realize a federated catalogue for data discovery from multiple EO data centers, the CEOS WGISS Integrated Catalog (CWIC) system was built. CWIC was initiated and supported by NASA, NOAA, and USGS as a contribution to CEOS to provide inventory search to WGISS agency catalog systems for EO data by distributing search requests to the appropriate server and sending search responses back to the requesting client.   The CWIC functionality is now supported within NASA/CMR for WCDA collections.

FedEO (Federated Earth Observation Gateway) provides a unique entry point to a growing number of scientific catalogues and services for, but not limited to, EO European and Canadian missions. FedEO is deployed with ESA (European Space Agency) infrastructure as a gateway to provide brokered discovery, access and ordering capability to European/Canadian EO missions data based on HMA (Heterogeneous Missions Accessibility) interfaces.

WGISS is now coordinating efforts to connect NASA and FedEO systems with IDN through a common registration of metadata records to seamlessly provide search results for relevant data sets regardless of which system is used to access the granule level data.

## 4.2  Search Concept and Design

A two-step collection/granule search process, which separates discovery of data collections from searching within relevant data collections to retrieve specific data granules, has been adopted to realize the integrated access to heterogeneous, autonomous data sources.

The WCDA system is an implementation of this two-step process. The IDN provides an OpenSearch front end to the collection search. The response from the collection search includes links to the OpenSearch Descriptor Documents (OSDD) at one of several Granule Gateways, providing search capability for granules at the relevant data providers. Current WCDA Granule Gateways include the CWIC/CMR and FedEO systems. Each of these systems provides access to different data archive systems using the same OpenSearch protocol. Spatial and temporal metadata are the only attributes guaranteed to be supported at all data providers. The CEOS OpenSearch Best Practice supports metadata attributes through the Parameter extensions, but they are not all universally supported across WCDA.

### 4.2.1 Collection Search Criteria

CEOS OpenSearch is used as the IDN's collections search implementation based on the OpenSearch 1.1 (Draft 5) specification and is compliant with the CEOS OpenSearch Best

Practices. The CEOS OpenSearch allows clients to formulate OpenSearch compliant queries against the IDN collections and specify the desired search results format as OpenSearch compliant ATOM or HTML. The IDN OpenSearch API has implementation the following search fields for users' queries:

- boundingBox
- campaign
- endTime
- geometry
- instrument
- keyword
- placeName
- processingLevel
- satellite
- sensor
- startTime
- uid

Also, client developers can query with specific tags: isCeos, isCwic, isGeoss, and isFedEO. Tagging allows arbitrary sets of collections to be grouped under a single namespace value. The sets of collections can be recalled later when searching by tag fields.

IDN query examples:

- GET the first 10 IDN collections with results in the ATOM format: https://cmr.earthdata.nasa.gov/opensearch/collections.atom?numberOfResults=10&clientId=wcdaOpenSearchDoc
- GET the first 10 IDN collections containing the GCMD instrument keyword MODIS with results in the ATOM output format: https://cmr.earthdata.nasa.gov/opensearch/collections.atom?instrument=MODIS&numberOfResults=10&clientId=wcdaOpenSearchDoc
- GET the first 10 CWIC collection IDN collections containing the GCMD instrument keyword MODIS with results in the HTML format: https://cmr.earthdata.nasa.gov/opensearch/collections?instrument=MODIS&isCwic=true&numberOfResults=10&clientId=wcdaOpenSearchDoc

### 4.2.2 Granule Search Criteria

Section 5 (below) gives complete details about the OpenSearch API for specifying search criteria. It is important to understand how the OSDD documents search parameters both for data collections and, separately, for granules.

The step 2 search, *i.e.*, granule search, supports constraints with spatial, temporal, and pagination. Spatial constraint is specified with a bounding box. Temporal filter can be specified with begin and end time in standard ISO 8601 timestamp format. See Section 5 for specific query parameters.

### 4.2.3 Current Data Partners

Current data partners providing granule search access through WGISS Connected Data Assets include systems from the following agencies: National Space and Aeronautics Administration (NASA), the European Space Agency (ESA), the National Centers for Environmental Information (NCEI) of National Oceanic and Atmospheric Administration (NOAA), the Group for High Resolution Sea Surface Temperature (GHRSST) of NOAA, United States Geological Survey (USGS), the National Institute for Space Research (INPE) of Brazil, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the Canada Centre for Mapping and Earth Observation (CCMEO), the Meteorological and Oceanographic Satellite Data Archival Centre (MOSDAC) of the Indian Space Research Organisation (ISRO), and the National Remote Sensing Center (NRSC) of ISRO and the National Remote Sensing Center of China (NRSCC).   WGISS maintains a current list of the data providers that comprise the WGISS Connected Data Assets on the WGISS website.

# 4.3 WGISS Architecture for Connected Data Assets

At its core, the systems present to End Users and Clients an OpenSearch server compliant with the CEOS OpenSearch Best Practice.   In this way, outside clients need to have no specific knowledge of the partner data systems and communicate only via OpenSearch.

# 4.4 Terms and Definitions

For the purposes of this document, the following terms and definitions apply:

**(1) catalog ID**

   Identifiers of data provider serving granule metadata

**(2) client**

   A software component that can invoke an operation from a server

**(3) CMR**

   The Common Metadata Repository (CMR) of the National Aeronautics and Space Administration (NASA) is a high-performance, high-quality, continuously evolving metadata system that catalogs Earth Science data and associated service metadata records.

**(4) collection**

A grouping of granules that all come from the same source, such as a modeling group or institution. Collections have information that is common across all the granules they "own" and a template for describing additional attributes not already part of the metadata model.

(5) **data clearinghouse**

The collection of institutions providing digital data, which can be searched through a single interface using a common metadata standard

(6) **dataset**

Has the same meaning as collection, see (4)

**(7) GCMD Keywords**

The Global Change Master Directory (GCMD) Keywords are a hierarchical set of controlled Earth Science vocabularies that help ensure Earth science data, services, and variables are described in a consistent and comprehensive manner and allow for the precise searching of metadata and subsequent retrieval of data, services, and variables.

**(8) granule**

The smallest aggregation of data that can be independently managed (described, inventoried, and retrieved). Granules have their own metadata model and support values associated with the additional attributes defined by the owning collection.

(9) **Granule Gateway**

A CEOS OpenSearch compliant server providing access to remote data partner inventory systems.

(10) **granule ID**

A character string that uniquely identifies a single granule to a granule gateway

(11) **identifier**

A character string that may be composed of numbers and characters that is exchanged between the client and the server with respect to a specific identity of a resource

(12) **IDN**

The CEOS International Directory Network (IDN) is a gateway to earth science data and services.

**(13) IDN dataset ID**

Unique dataset identifier in IDN, returned from the IDN in response to the OSDD request. This identifier is assigned by the IDN CMR database and may be referred to as the "conceptID" in CMR-specific discussions.

**(14) native ID**

A local dataset identifier used by WCDA partners to retrieve granule metadata through data provider API. This identifier is assigned by the data provider but may be the same as the IDN dataset ID.

**(15) operation**

The specification of a transformation or query that an object may be called to execute

**(16) OSDD**

OpenSearch Description Document

**(17) profile**

A set of one or more base standards and - where applicable - the identification of chosen clauses, classes, subsets, options and parameters of those base standards that are necessary for accomplishing a particular function

**(18) request**

The invocation of an operation by a client

**(19) response**

The result of an operation, returned from server to client

## 4.5  IDN Systems

The IDN, CMR OpenSearch (for IDN), and the GCMD Keyword have operational systems which end-users can access.

- IDN site is available to all users.
  Location: https://idn.ceos.org/
- OpenSearch API for IDN (via CMR) - production instance is available to all users.
  Location: https://cmr.earthdata.nasa.gov/opensearch/
- GCMD Keywords
  Location: https://earthdata.nasa.gov/earth-observation-data/find-data/idn/gcmd-keywords
- GCMD Keyword Viewer
  Location: https://gcmd.earthdata.nasa.gov/KeywordViewer

The IDN site search interface and the CMR OpenSearch production instances will provide access to all datasets which have been registered in the IDN. The Keyword Management System

production instance will provide access to all approved GCMD keywords registered by IDN providers.

## 4.6  CWIC/CMR Systems

Collection metadata for the former CWIC data partners within WCDA are available directly from CMR. The IDN response records contain direct links to the OSDD from the remote partner server, where possible. This can be used to construct a granule search directly from the partner.

## 4.7  Contact Information

All the documents and information about WCDA partners are available at WGISS Connected Data Assets page at

        https://ceos.org/ourwork/workinggroups/wgiss/access/

Any questions regarding to WCDA, please send the email to

        Access-SysTeam-Help@wgiss.ceos.org

# 5  OpenSearch Query Interface

## 5.1 Introduction

OpenSearch is a light-weight search specification[1] used by the WGISS Connected Data Assets to search and return metadata related to granule-level inventory data.   The OpenSearch API is not designed, nor is it used for returning observational data from the inventory systems, although the metadata returned should include links directly to data granules or to a data ordering system. CEOS OpenSearch is intended to take the end user as close to actual data as possible within the constraints of the data partner inventory systems and the limits of the OpenSearch protocol itself. The WCDA Granule Gateways have the task of returning valid responses in Atom[2] format as per CEOS OpenSearch OSDD requests. These provide URLs for submitting search requests directly to the Data Partner's inventory system for the requested dataset metadata.

## 5.2 Obtaining OpenSearch Description Document (OSDD)

OpenSearch Description Documents (OSDDs) provide necessary information for clients to programmatically construct valid search requests. Specifically, clients can acquire both cardinality and domain of request parameters based on query template. Dataset valids (*i.e.* spatial footprint and temporal extent) are also provided through the OSDD in both machine parsable and human readable formats. With dataset valids, clients can construct requests yielding more accurate results.

CEOS OpenSearch servers for CWIC/CMR provide both generic and dataset specific OSDDs. Clients can fetch a generic OSDD through the CMR OSDD endpoint.   The OSDD request must also include a client identifier string, as recommended by the CEOS OpenSearch Best Practices. Clients can also retrieve a dataset-specific OSDD through the OSDD endpoint by sending both client ID and dataset identifier (*i.e.* CMR conceptId). In a dataset-specific OSDD, domain is also provided for some parameters (*i.e. timeStart* and *timeEnd*) in addition to the request parameter syntax.

**Generic OSDD request URL example (CWIC/CMR):**

https://cmr.earthdata.nasa.gov/opensearch/collections/descriptor_document.xml?clientId=wgissOpenSearchDoc

---

[1]  OpenSearch specification version 1.1 (https://www.opensearch.org/docs/1.1)

[2]  Atom syndication format (https://tools.ietf.org/search/rfc4287)

**Generic OSDD request URL example (FedEO):**

https://fedeo.ceos.org/opensearch/description.xml

**Dataset specific OSDD request URL example (CWIC/CMR):**

https://cmr.earthdata.nasa.gov/opensearch/granules/descriptor_document.xml?collectionConceptId=C1235542031-USGS_LTA

**Dataset specific OSDD request URL example (FedEO):**

https://fedeo.ceos-org/opensearch/description.xml?parentIdentifier=EOP:ESA:FEDEO:COLLECTIONS:CryoSat.products

## 5.3  Search request

CEOS OpenSearch supports both searching for data collections through the IDN and for granules in a specific dataset at one of the data partners. It executes a collection or inventory search, as appropriate, and returns the matching results.

In order to initialize a valid request, clients will fill request parameters with proper values and set the dataset identifier. The template of the OpenSearch request is available under the <Url> element in OSDD. Both cardinality and domain of request parameters to be extracted from the OSDD are listed as follows:

| Request Parameter [a] | Description [b] | Value & Cardinality (M) = mandatory (O) = optional | Type [c] |
|---|---|---|---|
| datasetId | Dataset identifier | (M) <br> Allowed value is IDN dataset ID | lta[d]:datasetId |
| geoBox | Returned granules will have a spatial extent overlapping this bounding box | (O) <br> Supported formats are in W,S,E,N coordinate order: <br> W: WestBoundingLongitude <br> S: SouthBoundingLatitude <br> E: EastBoundingLongitude <br> N: NorthBoundingLatitude <br> All coordinates are in EPSG:4326 | geo[e]:box |

| timeStart | Returned granules will have a temporal extent containing this start time | (O) Supported formats are: 'yyyy-MM-dd', 'yyyy-MM-ddTHH:mm:ssZ' or 'yyyy-MM-dd HH:mm:ss' | time[f]:start |
|---|---|---|---|
| timeEnd | Returned granules will have a temporal extent containing this end time | (O) Supported formats are: 'yyyy-MM-dd', 'yyyy-MM-ddTHH:mm:ssZ' or 'yyyy-MM-dd HH:mm:ss' | time[f]:end |
| startIndex | Starting record number of the set of search results desired by the search client | (O) Allowed value is any integer equal and greater than '1'. | os[g]:startIndex |
| count | Number of search results per page desired by the search client | (O) Allowed value is any integer within the interval of [1,200]. | os[g]:count |
| clientId | The identifier of client | (O) Allowed value is any URL well-formed string representing client identifier. | esipdiscover[h]:clientId |

a: All request parameters are case sensitive

b: "Definition" represents the semantic meaning of request parameter.

c: "Type" represents request parameter type restricted by namespace.

d: /opensearch/extensions/1.0/

e: https://a9.com/-/opensearch/extensions/geo/1.0/

f: https://a9.com/-/opensearch/extensions/time/1.0/

g: https://a9.com/-/spec/opensearch/1.1/

h: https://commons.esipfed.org/ns/discovery/1.2/

Table 1 Table of CEOS OpenSearch request parameters

## 5.4 OpenSearch response

An OpenSearch response is an ATOM feed with zero or more ATOM entry elements. Each entry element represents the metadata of a single granule pertaining to the query submitted within a set of results also defined by the query.

An ATOM response is a single ATOM feed element containing the following,
1.  Information about the search conducted in terms of title, author and ID.

2. Information about the nature of the result set in terms of total number of results, number of results returned and how many the client asked for.
3. Navigation information for traversing that result set including links to the previous, next, first and last results in the set.
4. Zero or more entries pertaining to granule metadata matching the client query

An entry is an ATOM entry element containing the following,
1. Basic information about returned granule in terms of title, author, summary and ID.
2. Information about the spatial footprint of returned granule.
3. Information about the temporal extent of returned granule.
4. Collection of ATOM link elements containing endpoints of granule browse, order and metadata for the returned granule.

# 6  Partner Guidelines

## 6.1 Metadata & Semantic Mapping

### 6.1.1 HTTP/HTTPS access

Although WCDA systems may attempt to use any mechanism available for connecting to Data Partners' data management systems in order to access the available inventory search, there are a few specifics which make the process simpler and more robust. The use of HTTP/HTTPS for accessing the inventory search engine is strongly preferred. This is widely used already, as web browsers are nearly universal and provide an effective user interface for both human and automated access. The CEOS Best Practices document mandates the use of Atom XML for responses from the partner data systems. Other structured formats like comma-delimited tables or JSON may be acceptable.

### 6.1.2 Spatial search

All WCDA data partners are expected to support some level of spatial search since the inventory data are anticipated to have a spatial component. For each granule, a simple bounding box, with the bounding coordinates individually identified, is the minimum required, although more complex spatial footprint geometries are possible in the future. The number of spatial footprints in data partner's response is not limited to one.

It is desirable to have the API also support a dynamic call to return the limits of the spatial search, although not necessary. The presence of such a service can help clients avoid invalid or inappropriate search requests, such as those outside the spatial boundaries for specific data collections.

### 6.1.3 Temporal search

Like spatial search described in the previous section, all WCDA data partners are expected to support some level of temporal search since the inventory data are anticipated to have a temporal component. Simple temporal extent, with the start and end times individually identified is the

minimum required, although more complex temporal relations are anticipated in the future. It is best to support some minimal subset of the ISO 8601 time specification for syntax – YYYY-MM-DD, at least.

It is desirable to have the API also support a dynamic call to return the limits of the temporal extent search, although not necessary. The presence of such a service can help clients avoid invalid or inappropriate search requests, such as those outside the existing temporal extent for specific data collections.

## 6.1.4 Request for dataset by ID

Dataset ID (*cwic:datasetId* or *eo:parentIdentifier*) serves as a mandatory parameter in the WCDA OpenSearch query interface. Clients will query against data partner's search system, using the dataset ID, to retrieve the matching granule level metadata. From this point, all data partners are expected to support the query based on dataset identifier. In particular, a globally unique dataset identifier (*i.e.*, CMR conceptID or IDN *datasetID*) is recommended to be implemented as query parameter on data partner side. However, other unique dataset identifiers within data partner system are also acceptable.

## 6.1.5 Request for granule by ID

CEOS OpenSearch supports searching for a single granule by granule ID. Clients will query against data partner's search system with the granule ID in order to retrieve the matching granule level metadata. From this point, all data partners are recommended to support the query based on granule identifier.

## 6.1.6 Unique granule IDs

Each data granule returned in a search response should have an identifier associated with it which is unique within the dataset. It is important that the search response include the unique identifier for each granule so that the full data on individual granules may be retrieved without re-executing a (potentially time-consuming) search.

## 6.1.7 Record counts

As part of the search response from the inventory system, it is highly desirable to have the total count of matching granules returned, even if the metadata for the granules is not contained in the

search response. This parameter, coupled with the ability to specify the starting record number and number of desired records from the inventory system, will allow clients to implement results paging and reducing the load on both the WCDA systems and on the data partners.

## 6.1.8 Pagination

Pagination is supported (*i.e.,* search by specifying start index and counts per page) in the WCDA OpenSearch query interface. The search based on pagination parameter (*e.g.*, start index and items per page) or cursor parameter (*e.g.*, start index and items per page) is expected to be implemented by the Data Partner.

## 6.1.9 Other search parameters

The Data Partner inventory systems may support search on additional parameters like LANDSAT path/row.

## 6.2 Interaction & Services Model

## 6.2.1 Unique granule ID

As described above, each data granule should have a unique identifier which a) is passed back to the client as part of the search response and b) can be used as a key with which to retrieve that specific granule. The OSDD will provide the association between the IDN identifier and the correct dataset and data center.

## 6.2.2 Contact information

The CEOS OpenSearch OSDD generally includes contact information for each collection. These are usually the same for all data granules, and frequently the same for all datasets within a single data center.

There is no need for this information to be returned with each search response or each data granule, although it might be.

## 6.2.3 Browse URL

If browse images of the data granule are available, a valid URL to display the browse image should be included in the search response for each granule so that the client can display it as a link.

## 6.2.4 Order URL

The CEOS Best Practices recommends that, when the granule data can be downloaded from the data center directly, a valid URL to retrieve the data be included in the search response for the granule.

Alternatively, the search response may contain a URL directing the user to a web site for ordering the data if this is the only option permitted by the data center. This is often necessary even for freely available data if, for example, data center policies require user registration before downloads are made available.

## 6.2.5 Pagination support

Pagination parameters *startIndex* and count are stipulated by CEOS OpenSearch Best Practices and are expected to be supported for each Data Partner.

## 6.2.6 Spatial search

As stated above, a Data Partner is expected to support spatial search and include proper spatial component in each granule response. Native spatial components (one or many) from data partner may be included, but responses should also minimally *georss:box* or *georss:polygon*, which are permitted by CEOS OpenSearch Best Practices. Besides that, a minimum bounding rectangle (MBR) should also be included in CEOS OpenSearch response as some clients may not support more complex geometric structures.

## 6.2.7 Temporal search

As stated above, a Data Partner is expected to support temporal search and include standard (ISO 8601 compliant) temporal elements in each granule response. Temporal elements with native

format from data partner are converted syntactically and semantically to single temporal element (*i.e. dc:date*) which are stipulated by CEOS OpenSearch Best Practices.

## 6.3 Error Handling

Error handling in CEOS OpenSearch is based on standard HTTP/HTTPS status codes. Effort has also been dedicated to provide sensible error messages in addition to the generic HTTP/HTTPS status code to the end user or client.   In order to support this eventuality, it would be useful for the inventory search system to attempt to return sensible and relevant HTTP/HTTPS status codes (where applicable) if something goes wrong with the search or, perhaps even better, a small, descriptive response document (in XML or JSON or whatever the default format might be) providing error codes and error text. In this way, the clients can distinguish the type of error arising at the inventory system from those arising elsewhere and take appropriate action. See the WGISS CDA Error Handling Guide for additional details.

Useful status and error messages help the clients manage user sessions effectively. Any limitations on submitted search requests to the inventory systems should be noted in the response (*e.g.,* "too many records requested", "search timed out") so that predictable error-handling can be managed by the client. It is especially important to have a clear error response if there is any limitation on spatial and temporal values, which are two of the most common search specifications used against individual data collections.

Search timeouts are particularly difficult to identify and manage because they are unpredictable. In many cases, the remote server loses is unable to pass a status code to the client because the connection simply disappears. The underlying search engine or database might simply take too long to build a very large set of results, resulting in the connection from the client timing out. Limitations on spatial or temporal ranges might be imposed by the search engine or database but there is no current mechanism to make these limitations known in the OpenSearch OSDD. Best practices for handling such unpredictable situations are under discussion by the WCDA System Level team.

The CEOS OpenSearch Best Practices (CEOS-OPENSEARCH-BP-V1.3) recommends the use of HTTP/HTTPS 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.