



EOSDIS

NASA'S EARTH OBSERVING SYSTEM
DATA AND INFORMATION SYSTEM

Discovering NASA's Cloud Holdings

Winter ESIP 2023

Mark Reese

Search & Discovery Product Manager

mreese@element84.com

This work was supported by NASA/GSFC under Raytheon Technologies contract number 80GSFC21CA001.
This document does not contain technology or Technical Data controlled under either the U.S. International Traffic
in Arms Regulations or the U.S. Export Administration Regulations.

Discovering Cloud Holdings

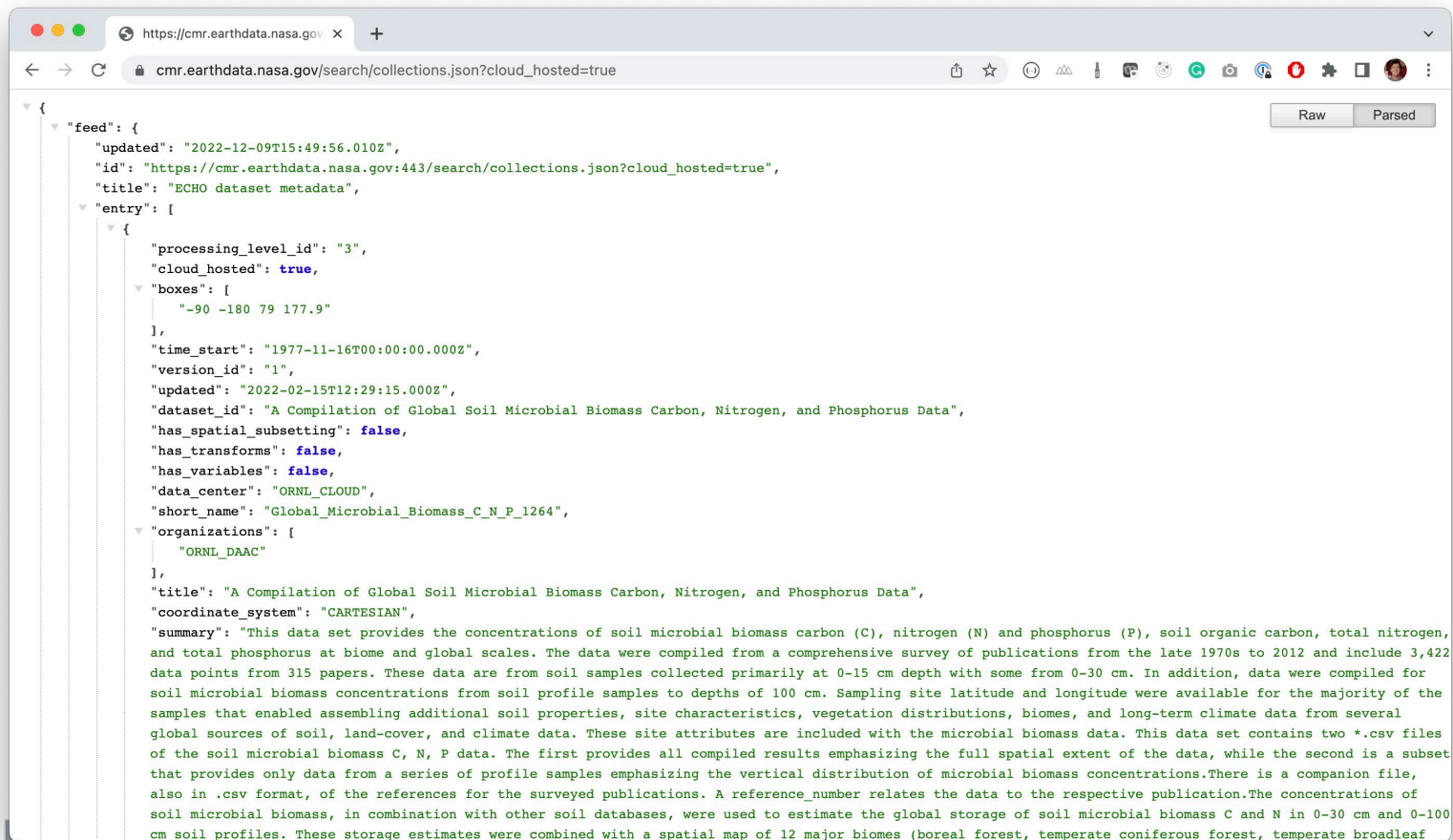
- CMR API
- GraphQL
- CMR STAC
- Earthdata Search

CMR API

- Pros
 - Most features
 - Multiple response format options
 - Will have latest capabilities first
- Cons
 - Can be overwhelming
 - Requires the most technical knowledge of all the options
 - Responses will likely include a lot of data you don't need

CMR API

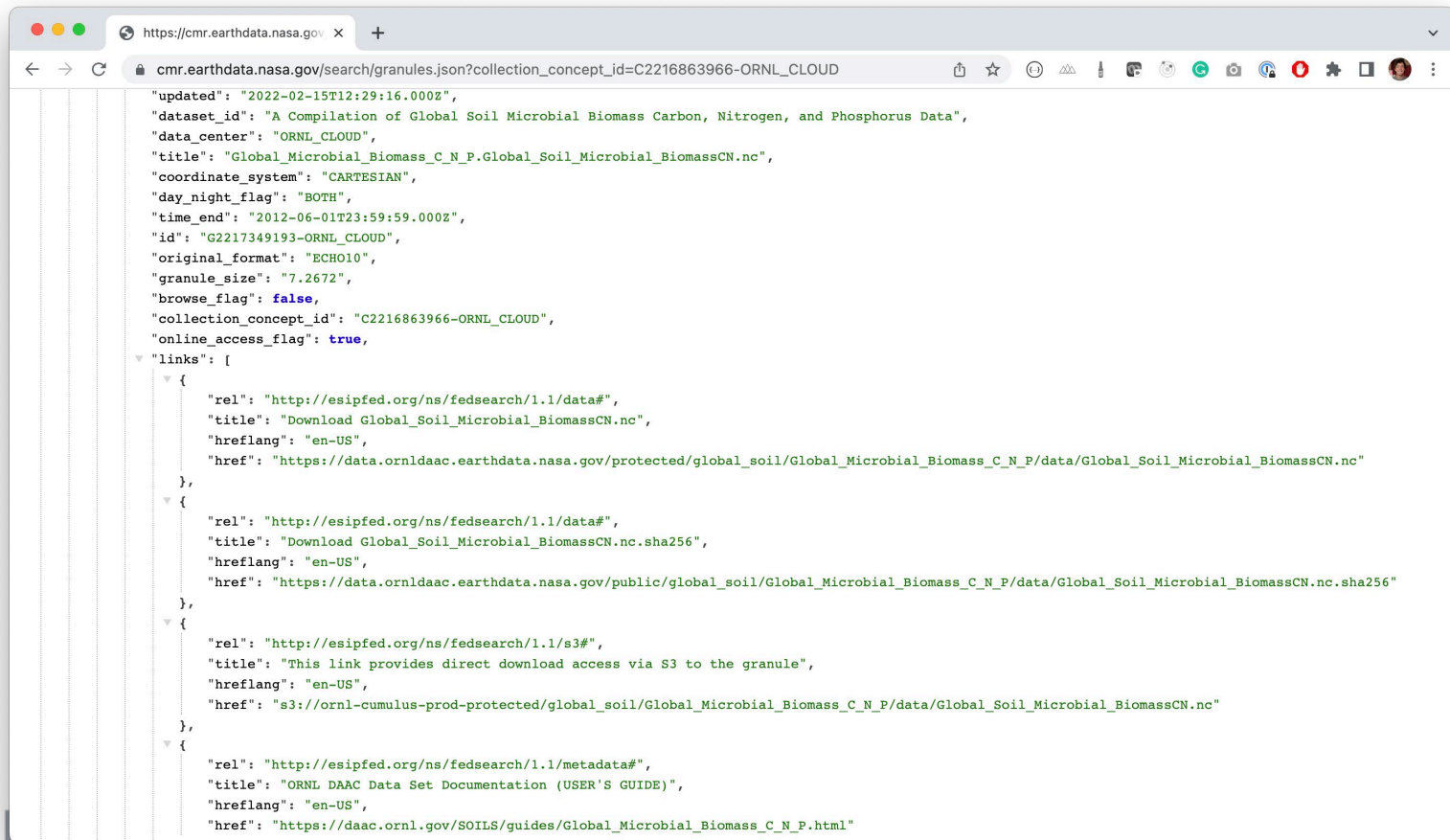
https://cmr.earthdata.nasa.gov/search/collections.json?cloud_hosted=true



```
{
  "feed": {
    "updated": "2022-12-09T15:49:56.010Z",
    "id": "https://cmr.earthdata.nasa.gov:443/search/collections.json?cloud_hosted=true",
    "title": "ECHO dataset metadata",
    "entry": [
      {
        "processing_level_id": "3",
        "cloud_hosted": true,
        "boxes": [
          "-90 -180 79 177.9"
        ],
        "time_start": "1977-11-16T00:00:00.000Z",
        "version_id": "1",
        "updated": "2022-02-15T12:29:15.000Z",
        "dataset_id": "A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data",
        "has_spatial_subsetting": false,
        "has_transforms": false,
        "has_variables": false,
        "data_center": "ORNL_CLOUD",
        "short_name": "Global_Microbial_Biomass_C_N_P_1264",
        "organizations": [
          "ORNL_DAAC"
        ],
        "title": "A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data",
        "coordinate_system": "CARTESIAN",
        "summary": "This data set provides the concentrations of soil microbial biomass carbon (C), nitrogen (N) and phosphorus (P), soil organic carbon, total nitrogen, and total phosphorus at biome and global scales. The data were compiled from a comprehensive survey of publications from the late 1970s to 2012 and include 3,422 data points from 315 papers. These data are from soil samples collected primarily at 0-15 cm depth with some from 0-30 cm. In addition, data were compiled for soil microbial biomass concentrations from soil profile samples to depths of 100 cm. Sampling site latitude and longitude were available for the majority of the samples that enabled assembling additional soil properties, site characteristics, vegetation distributions, biomes, and long-term climate data from several global sources of soil, land-cover, and climate data. These site attributes are included with the microbial biomass data. This data set contains two *.csv files of the soil microbial biomass C, N, P data. The first provides all compiled results emphasizing the full spatial extent of the data, while the second is a subset that provides only data from a series of profile samples emphasizing the vertical distribution of microbial biomass concentrations. There is a companion file, also in .csv format, of the references for the surveyed publications. A reference_number relates the data to the respective publication. The concentrations of soil microbial biomass, in combination with other soil databases, were used to estimate the global storage of soil microbial biomass C and N in 0-30 cm and 0-100 cm soil profiles. These storage estimates were combined with a spatial map of 12 major biomes (boreal forest, temperate coniferous forest, temperate broadleaf
```

CMR API

https://cmr.earthdata.nasa.gov/search/granules.json?collection_concept_id=C2216863966-ORNL_CLOUD



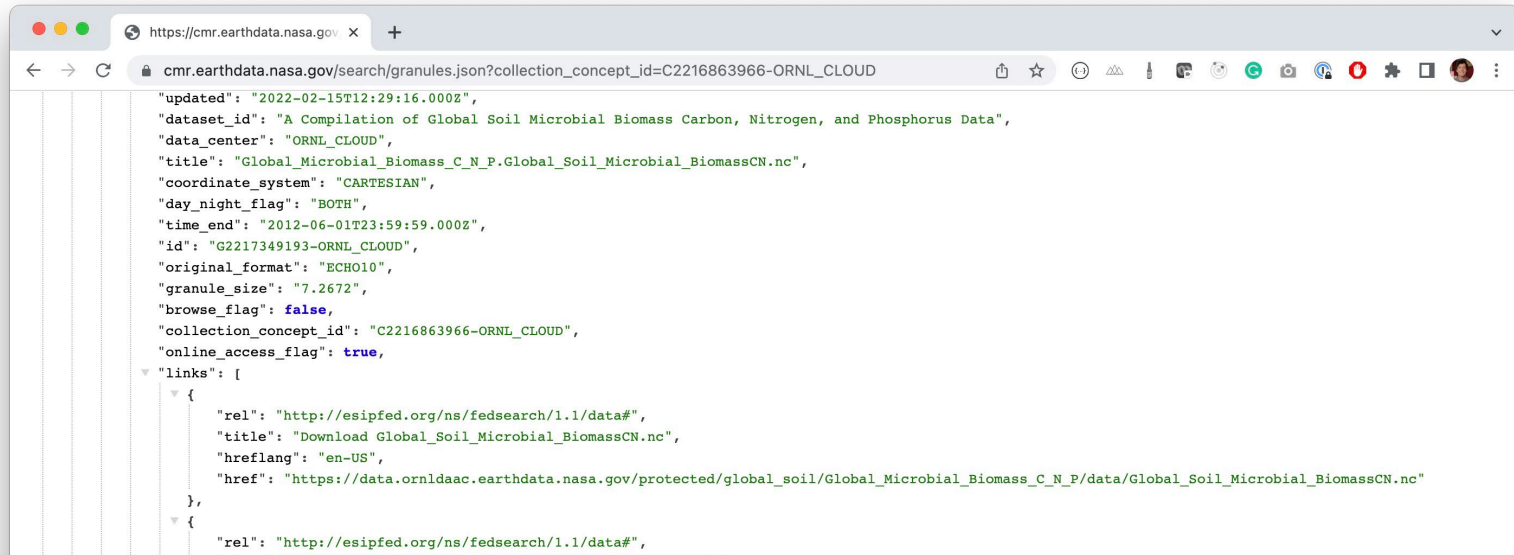
The screenshot shows a web browser window with the URL https://cmr.earthdata.nasa.gov/search/granules.json?collection_concept_id=C2216863966-ORNL_CLOUD. The browser displays a JSON response from the CMR API. The JSON object contains metadata about a dataset and a list of links for downloading the data.

```
{
  "updated": "2022-02-15T12:29:16.000Z",
  "dataset_id": "A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data",
  "data_center": "ORNL_CLOUD",
  "title": "Global_Microbial_Biomass_C_N_P.Global_Soil_Microbial_BiomassCN.nc",
  "coordinate_system": "CARTESIAN",
  "day_night_flag": "BOTH",
  "time_end": "2012-06-01T23:59:59.000Z",
  "id": "G2217349193-ORNL_CLOUD",
  "original_format": "ECHO10",
  "granule_size": "7.2672",
  "browse_flag": false,
  "collection_concept_id": "C2216863966-ORNL_CLOUD",
  "online_access_flag": true,
  "links": [
    {
      "rel": "http://esipfed.org/ns/fedsearch/1.1/data#",
      "title": "Download Global_Soil_Microbial_BiomassCN.nc",
      "hreflang": "en-US",
      "href": "https://data.ornldaac.earthdata.nasa.gov/protected/global_soil/Global_Microbial_Biomass_C_N_P/data/Global_Soil_Microbial_BiomassCN.nc"
    },
    {
      "rel": "http://esipfed.org/ns/fedsearch/1.1/data#",
      "title": "Download Global_Soil_Microbial_BiomassCN.nc.sha256",
      "hreflang": "en-US",
      "href": "https://data.ornldaac.earthdata.nasa.gov/public/global_soil/Global_Microbial_Biomass_C_N_P/data/Global_Soil_Microbial_BiomassCN.nc.sha256"
    },
    {
      "rel": "http://esipfed.org/ns/fedsearch/1.1/s3#",
      "title": "This link provides direct download access via S3 to the granule",
      "hreflang": "en-US",
      "href": "s3://ornl-cumulus-prod-protected/global_soil/Global_Microbial_Biomass_C_N_P/data/Global_Soil_Microbial_BiomassCN.nc"
    },
    {
      "rel": "http://esipfed.org/ns/fedsearch/1.1/metadata#",
      "title": "ORNL DAAC Data Set Documentation (USER'S GUIDE)",
      "hreflang": "en-US",
      "href": "https://daac.ornl.gov/SOILS/guides/Global_Microbial_Biomass_C_N_P.html"
    }
  ]
}
```



CMR API

https://cmr.earthdata.nasa.gov/search/granules.json?collection_concept_id=C2216863966-ORNL_CLOUD



A screenshot of a web browser displaying the JSON response from the CMR API. The browser's address bar shows the URL: `https://cmr.earthdata.nasa.gov/search/granules.json?collection_concept_id=C2216863966-ORNL_CLOUD`. The JSON response is as follows:

```
{
  "updated": "2022-02-15T12:29:16.000Z",
  "dataset_id": "A Compilation of Global Soil Microbial Biomass Carbon, Nitrogen, and Phosphorus Data",
  "data_center": "ORNL_CLOUD",
  "title": "Global_Microbial_Biomass_C_N_P.Global_Soil_Microbial_BiomassCN.nc",
  "coordinate_system": "CARTESIAN",
  "day_night_flag": "BOTH",
  "time_end": "2012-06-01T23:59:59.000Z",
  "id": "G2217349193-ORNL_CLOUD",
  "original_format": "ECHO10",
  "granule_size": "7.2672",
  "browse_flag": false,
  "collection_concept_id": "C2216863966-ORNL_CLOUD",
  "online_access_flag": true,
  "links": [
    {
      "rel": "http://esipfed.org/ns/fedsearch/1.1/data#",
      "title": "Download Global_Soil_Microbial_BiomassCN.nc",
      "hreflang": "en-US",
      "href": "https://data.ornldaac.earthdata.nasa.gov/protected/global_soil/Global_Microbial_Biomass_C_N_P/data/Global_Soil_Microbial_BiomassCN.nc"
    },
    {
      "rel": "http://esipfed.org/ns/fedsearch/1.1/data#",
```

```
{
  "rel": "http://esipfed.org/ns/fedsearch/1.1/s3#",
  "title": "This link provides direct download access via S3 to the granule",
  "hreflang": "en-US",
  "href": "s3://ornl-cumulus-prod-protected/global_soil/Global_Microbial_Biomass_C_N_P/data/Global_Soil_Microbial_BiomassCN.nc"
},
```

```
{
  "rel": "http://esipfed.org/ns/fedsearch/1.1/metadata#",
  "title": "ORNL DAAC Data Set Documentation (USER'S GUIDE)",
  "hreflang": "en-US",
  "href": "https://daac.ornl.gov/SOILS/guides/Global_Microbial_Biomass_C_N_P.html"
}
```



CMR GraphQL

- Pros
 - Supports *most* CMR features
 - Simplified UX
 - Playground for experimenting with requests that contains documentation
 - Ask for specific fields and only get what you request
- Cons
 - Will slightly lag behind CMR features
 - Performance can be slightly slower than hitting CMR directly

CMR GraphQL

<https://graphql.earthdata.nasa.gov/api>

The screenshot displays the Apollo GraphQL Studio interface in a web browser. The address bar shows the URL `studio.apollographql.com/sandbox/explorer`. The interface is divided into several panels:

- Documentation:** On the left, it shows the schema path `Root > Query > collections > (params)` and the type `params: CollectionsInput`. Below this, the "Input Fields" section lists various parameters like `boundingBox`, `circle`, `cloudHosted` (which is selected), `collectionDataType`, `conceptId`, `cursor`, `dataCenter`, `dataCenterH`, `dataCenters`, `facetsSize`, `granuleDataFormat`, `granuleDataFormatH`, `hasGranules`, and `hasGranulesOrCwic`.
- Operation:** The central panel shows a GraphQL query:

```
1 query($collectionParams: CollectionsInput) { ...
2   collections(params: $collectionParams) {
3     count
4     items {
5       datasetId
6     }
7   }
8 }
```

A "Run" button is visible next to the query.
- Variables:** Below the query, the "Variables" section shows the input variables in JSON:

```
1 {
2   "collectionParams": {
3     "cloudHosted": true
4   }
5 }
```
- Response:** On the right, the JSON response is displayed with a status of 200, a response time of 843ms, and a size of 1.9KB. The response structure is:

```
{
  "data": {
    "collections": {
      "count": 1951,
      "items": [
        {
          "datasetId": "A Compilation of Global
Soil Microbial Biomass Carbon, Nitrogen, and
Phosphorus Data"
        },
        {
          "datasetId": "A Concise Experiment Plan
for the Arctic-Boreal Vulnerability Experiment"
        },
        {
          "datasetId": "A Global Database of Gas
Fluxes from Soils after Rewetting or Thawing,
Version 1.0"
        },
        {
          "datasetId": "A Global Database of Soil
Phosphorus Compiled from Studies Using Hedley
Fractionation"
        },
        {
          "datasetId": "A Global Database of Soil
Respiration Data, Version 5.0"
        }
      ]
    }
  }
}
```

CMR GraphQL

<https://graphql.earthdata.nasa.gov/api>

The screenshot displays the Apollo GraphQL Studio interface in a web browser. The address bar shows the URL `studio.apollographql.com/sandbox/explorer`. The interface is divided into several panels:

- Documentation:** On the left, it shows the schema path `Root > Query > collections > (params)` and the `params: CollectionsInput` type. Below this, the **Input Fields** section lists various query parameters like `boundingBox`, `circle`, `cloudHosted`, `collectionDataType`, `conceptId`, `cursor`, `dataCenter`, `dataCenterH`, `dataCenters`, `facetsSize`, `granuleDataFormat`, `granuleDataFormatH`, `hasGranules`, and `hasGranulesOrCwic`.
- Operation:** The central panel shows a GraphQL query:

```
1 query($collectionParams: CollectionsInput) { ...
2   collections(params: $collectionParams) {
3     count
4     items {
5       datasetId
6       granules {
7         items {
8           links
9         }
10      }
11    }
12  }
13 }
```
- Variables:** Below the query, the **Variables** section shows the JSON input:

```
1 {
2   "collectionParams": {
3     "cloudHosted": true
4     "conceptId": "C1940473819-POCLOUD"
5   }
6 }
```
- Response:** The right panel displays the JSON response from the API. It includes a `data` object with `collections` containing `count` and `items`. The `items` array contains a single object with `datasetId`, `granules`, and `links`. The `links` array contains a single object with `rel`, `title`, `hreflang`, and `href`. The `href` value is a long URL: `http://esipfed.org/ns/fedsearch/1.1/s3#`.

CMR GraphQL

<https://graphql.earthdata.nasa.gov/api>

Operation



Run

```
1 query($collectionParams: CollectionsInput) { ...
2   collections(params: $collectionParams) {
3     count
4     items {
5       datasetId
6       granules {
7         items {
8           links
9         }
10      }
11    }
12  }
13 }
```

CMR GraphQL

<https://graphql.earthdata.nasa.gov/api>

```
10  |   |   |  
11  |   |   }  
12  |   }  
13  }
```

Variables Headers

```
1  {  
2    "collectionParams": {  
3      "cloudHosted": true  
4      "conceptId": "C1940473819-POCLOUD"  
5    }  
6  }
```

JSON

CMR GraphQL

<https://graphql.earthdata.nasa.gov/api>

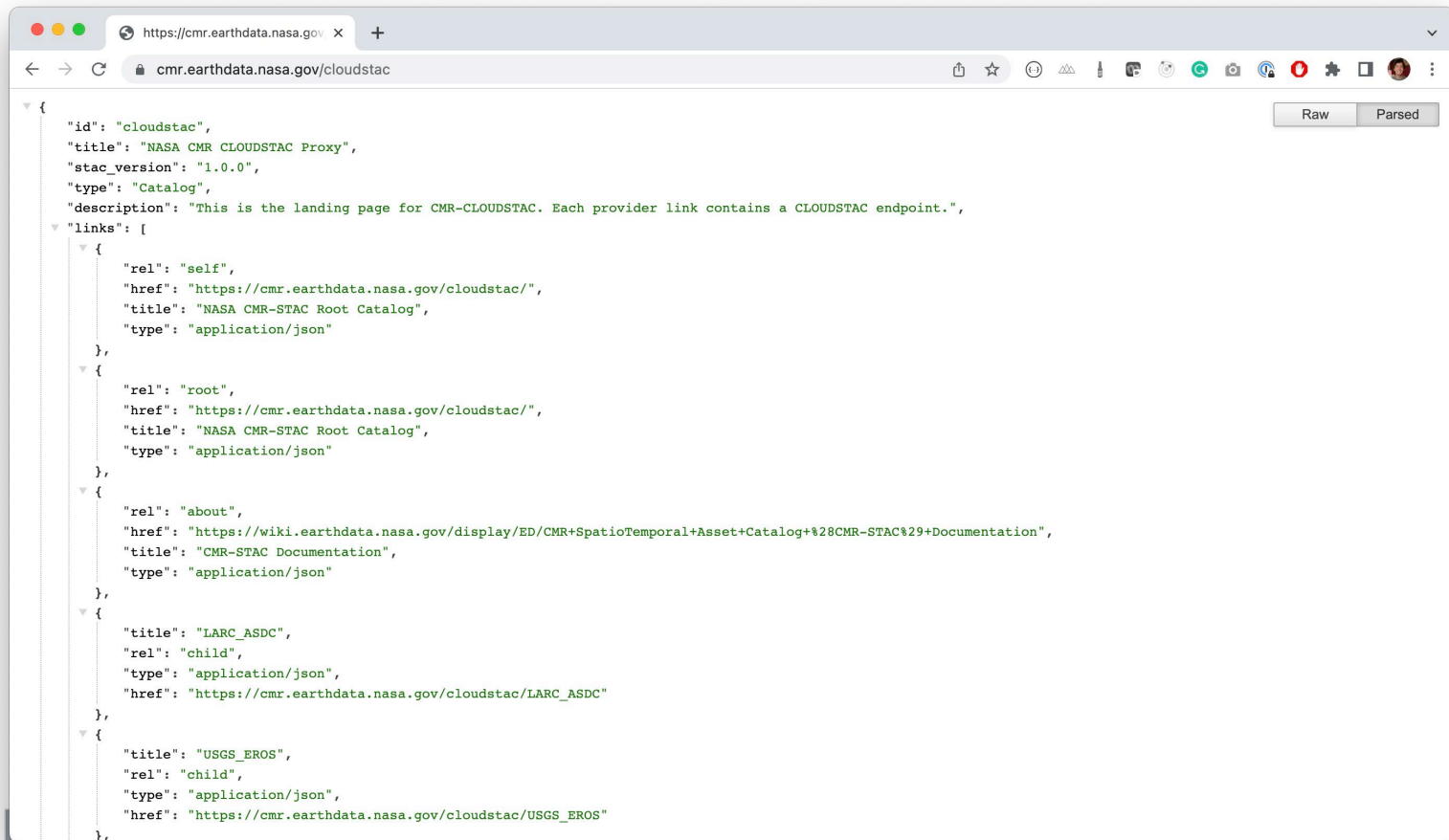
```
{
  "data": {
    "collections": {
      "count": 1,
      "items": [
        {
          "datasetId": "GHRSSST Level 2P Global
Sea Surface Skin Temperature from the Moderate
Resolution Imaging Spectroradiometer (MODIS) on
the NASA Aqua satellite (GDS2)",
          "granules": {
            "items": [
              {
                "links": [
                  {
                    "rel": "http://esipfed.org/ns/
fedsearch/1.1/s3#",
                    "title": "This link provides
direct download access via S3 to the granule.",
                    "hreflang": "en-US",
                    "href": "s3://
podaac-ops-cumulus-protected/
MODIS_A-JPL-L2P-v2019.0/
20020704000015-JPL-L2P_GHRSSST-SSTskin-MODIS_A-N-v0
"
```

CMR STAC

- Pros
 - Complies with STAC v1.0
 - There is a specific Cloud STAC catalog to simplify searching/browsing
- Cons
 - Does not support advanced granule searching like CMR
 - As of right now, S3 URLs are not surfaced in STAC format

CMR STAC

<https://cmr.earthdata.nasa.gov/cloudstac>



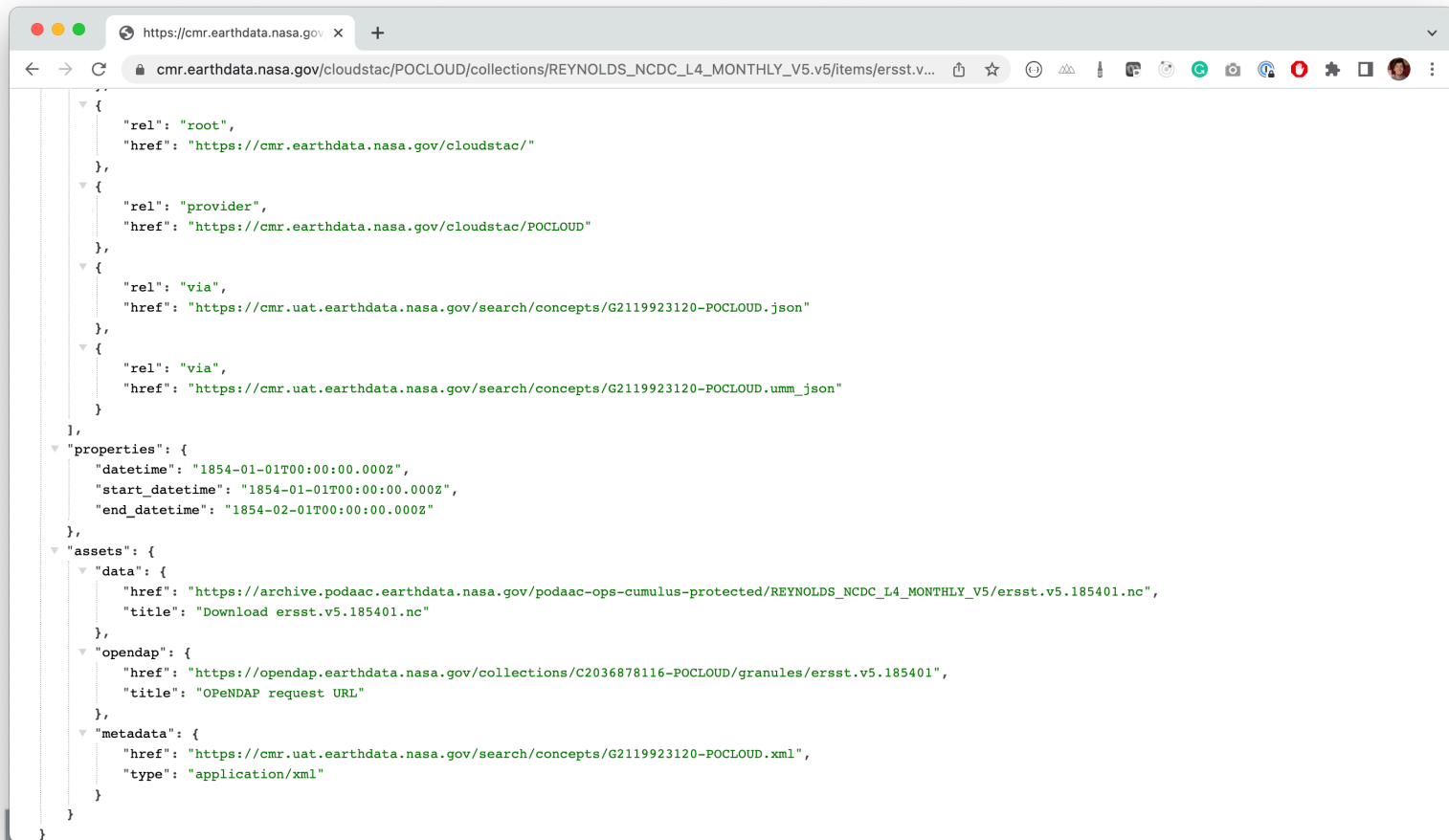
The screenshot shows a web browser window with the address bar displaying <https://cmr.earthdata.nasa.gov/cloudstac>. The page content is a JSON document representing a STAC Catalog. The JSON is displayed in a code editor with syntax highlighting. On the right side of the editor, there are two buttons: "Raw" and "Parsed". The JSON structure includes a root object with fields for id, title, stac_version, type, description, and a links array. The links array contains five objects representing different links: self, root, about, LARC_ASDC, and USGS_EROS.

```
{
  "id": "cloudstac",
  "title": "NASA CMR CLOUDSTAC Proxy",
  "stac_version": "1.0.0",
  "type": "Catalog",
  "description": "This is the landing page for CMR-CLOUDSTAC. Each provider link contains a CLOUDSTAC endpoint.",
  "links": [
    {
      "rel": "self",
      "href": "https://cmr.earthdata.nasa.gov/cloudstac/",
      "title": "NASA CMR-STAC Root Catalog",
      "type": "application/json"
    },
    {
      "rel": "root",
      "href": "https://cmr.earthdata.nasa.gov/cloudstac/",
      "title": "NASA CMR-STAC Root Catalog",
      "type": "application/json"
    },
    {
      "rel": "about",
      "href": "https://wiki.earthdata.nasa.gov/display/ED/CMR+SpatioTemporal+Asset+Catalog+%28CMR-STAC%29+Documentation",
      "title": "CMR-STAC Documentation",
      "type": "application/json"
    },
    {
      "title": "LARC_ASDC",
      "rel": "child",
      "type": "application/json",
      "href": "https://cmr.earthdata.nasa.gov/cloudstac/LARC_ASDC"
    },
    {
      "title": "USGS_EROS",
      "rel": "child",
      "type": "application/json",
      "href": "https://cmr.earthdata.nasa.gov/cloudstac/USGS_EROS"
    }
  ]
}
```



CMR STAC

<https://cmr.earthdata.nasa.gov/cloudstac>



The screenshot shows a web browser window with the address bar displaying `https://cmr.earthdata.nasa.gov/cloudstac/POCLOUD/collections/REYNOLDS_NCDC_L4_MONTHLY_V5.v5/items/ersst.v5.185401.nc`. The main content area displays a JSON object representing a STAC item. The JSON structure includes a `links` array with four entries: `root`, `provider`, `via` (two instances), and `properties`. The `properties` object contains `datetime`, `start_datetime`, and `end_datetime`. The `assets` array contains three entries: `data`, `opendap`, and `metadata`. Each asset entry includes a `href` and a `title`.

```
{
  "links": [
    {
      "rel": "root",
      "href": "https://cmr.earthdata.nasa.gov/cloudstac/"
    },
    {
      "rel": "provider",
      "href": "https://cmr.earthdata.nasa.gov/cloudstac/POCLOUD"
    },
    {
      "rel": "via",
      "href": "https://cmr.uat.earthdata.nasa.gov/search/concepts/G2119923120-POCLOUD.json"
    },
    {
      "rel": "via",
      "href": "https://cmr.uat.earthdata.nasa.gov/search/concepts/G2119923120-POCLOUD.umm_json"
    }
  ],
  "properties": {
    "datetime": "1854-01-01T00:00:00.000Z",
    "start_datetime": "1854-01-01T00:00:00.000Z",
    "end_datetime": "1854-02-01T00:00:00.000Z"
  },
  "assets": [
    {
      "data": {
        "href": "https://archive.podaac.earthdata.nasa.gov/podaac-ops-cumulus-protected/REYNOLDS_NCDC_L4_MONTHLY_V5/ersst.v5.185401.nc",
        "title": "Download ersst.v5.185401.nc"
      },
      "opendap": {
        "href": "https://opendap.earthdata.nasa.gov/collections/C2036878116-POCLOUD/granules/ersst.v5.185401",
        "title": "OPeNDAP request URL"
      },
      "metadata": {
        "href": "https://cmr.uat.earthdata.nasa.gov/search/concepts/G2119923120-POCLOUD.xml",
        "type": "application/xml"
      }
    }
  ]
}
```



CMR STAC

<https://cmr.earthdata.nasa.gov/cloudstac>



Earthdata Search

- Pros
 - Browser Based
 - 1 click to cloud-hosted data
 - Tips & tricks embedded throughout the user journey
 - Integration with enterprise services
- Cons
 - Browser Based
 - Development-heavy

Earthdata Search

<https://search.earthdata.nasa.gov>

The screenshot displays the Earthdata Search interface. The top navigation bar includes the NASA logo, 'EARTHDATA', and a search bar. Below this, the 'EARTHDATA SEARCH' header is visible. The main content area shows search results for 'GHRSSST Level 2P Global Sea Surface Skin Temperature from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the NASA Aqua satellite (GDS2)'. The results are displayed as a grid of granules, each with a thumbnail image and a table of start and end times. The left sidebar contains filters for Spatial (Polygon), Temporal (Start/Stop), and Filter Granules. The right sidebar shows a map of the Mediterranean region with a green polygon indicating the search area. The bottom of the page features a timeline for the month of July 2023, with a green bar representing the search results. The footer includes the version number 'v1.183.0', NASA Official: Stephen Berrick, and links to FOIA, NASA Privacy Policy, and USA.gov. A link to 'Earthdata Access: A Section 508 accessible alternative' is also present.

Search Results (143 Collections)

GHRSSST Level 2P Global Sea Surface Skin Temperature from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the NASA Aqua satellite (GDS2)

Showing 20 of 137 matching granules

Granule ID	START	END
20220730125501-JPL-L2P_GHRSSST-SSTskin-MODIS_A-D-v02.0-fv01.0	2022-07-30 12:55:01	2022-07-30 12:59:59
20220730111501-JPL-L2P_GHRSSST-SSTskin-MODIS_A-D-v02.0-fv01.0	2022-07-30 11:15:01	2022-07-30 11:19:59
20220730015001-JPL-L2P_GHRSSST-SSTskin-MODIS_A-N-v02.0-fv01.0	2022-07-30 01:50:01	2022-07-30 01:54:59
20220730014501-JPL-L2P_GHRSSST-SSTskin-MODIS_A-N-v02.0-fv01.0	2022-07-30 01:45:01	2022-07-30 01:50:00

Subscriptions Remove Download All 137

MONTH GHRSSST Level 2P Global Sea Surface Skin Te... Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun

v1.183.0 · NASA Official: Stephen Berrick · FOIA · NASA Privacy Policy · USA.gov

Earthdata Access: A Section 508 accessible alternative



Earthdata Search

<https://search.earthdata.nasa.gov>

The screenshot displays the Earthdata Search interface in a web browser. The search results for "GHRSSST Level 2P Global Sea Surface Skin Temperature from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the NASA Aqua satellite (GDS2)" are shown. The results list 137 matching granules, with two examples visible: 20220730125501-JPL-L2P_GHRSSST-SSTskin-MODIS_A-D-v02.0-fv01.0 and 20220730111501-JPL-L2P_GHRSSST-SSTskin-MODIS_A-D-v02.0-fv01.0. Each granule entry includes a thumbnail image, a "START" time, and an "END" time. A "Download Files" button is visible, and a pop-up window shows the "AWS S3 Access" information, including the region "us-west-2", the bucket/object prefix "podaac-ops-cumulus-public/MODIS_A-JPL-L2P-v2019.0/", and the AWS S3 credentials "Get AWS S3 Credentials" and "View Documentation". The interface also features a sidebar with filters for Spatial (Polygon), Temporal (Start/Stop), and Granule Search (Granule ID(s)). A map on the right shows the search area over the Mediterranean Sea and surrounding landmasses. The bottom of the page includes a timeline for the month of July 2023, a version number "v1.183.0", and links to NASA Official, FOIA, NASA Privacy Policy, and USA.gov.



Earthdata Search

<https://search.earthdata.nasa.gov>

The screenshot shows a web browser window displaying the 'Download Status' page for a specific dataset. The browser's address bar shows the URL `search.earthdata.nasa.gov/downloads/8436433946`. The page header includes the 'EARTHDATA SEARCH' logo and navigation links like 'Find a DAAC', 'Back to Project', and 'Mark'. The main heading is 'Download Status', followed by a note that the page updates automatically and provides a link to the 'Download Status and History' page.

The dataset title is 'GHR SST Level 2P Global Sea Surface Skin Temperature from the Moderate Resolution Imaging Spectroradiometer (MODIS) on...'. A table shows the status as 'Complete (100%)', the access method as 'Download', and the number of granules as '137 Granules'. Below this, a message states: 'Download your data directly from the links below, or use the provided download script.'

There are four tabs: 'Download Files', 'AWS S3 Access' (which is selected), 'Download Script', and 'Browse Imagery'. Under the 'AWS S3 Access' tab, it says 'Direct cloud access for this collection is available in the us-west-2 region in AWS S3.' A table provides the following information:

Region	Bucket/Object Prefix	AWS S3 Credentials
us-west-2	podaac-ops-cumulus-public/MODIS_A-JPL-L2P-v2019.0/	Get AWS S3 Credentials View Documentation

Below the table, it says 'Retrieved 137 objects for 137 granules'. A list of S3 object URLs is displayed, with buttons for 'Copy', 'Save', and 'Expand' at the top right of the list.

At the bottom of the page, there is a footer with the version 'v1.183.0', NASA Official: Stephen Berrick, and links to FOIA, NASA Privacy Policy, and USA.gov. On the right side of the footer, it says 'Earthdata Access: A Section 508 accessible alternative'.



This work was supported by NASA/GSFC under
Raytheon Technologies contract number
80GSFC21CA001.