



STAC & Federated Search

Doug Newman
NASA ESDIS Project
Goddard Space Flight Center

Summer ESIP
July 22nd 2024



Agenda

- CEOS, WGISS & CWIC
- Federated search based on two step
- OpenSearch
- STAC as the backbone of two-step federated search



The need for federated search

The Committee on Earth Observation Satellites (CEOS) ensures international coordination of civil space-based Earth observation programs and promotes exchange of data to optimize societal benefit and inform decision making for securing a prosperous and sustainable future for humankind.

The Working Group on Information Systems and Services (WGISS) is a CEOS working group that promotes collaboration in the development of systems and services that manage and supply observation data [...] creates and demonstrates prototypes supporting CEOS and Group on Earth Observation (GEO) requirements.

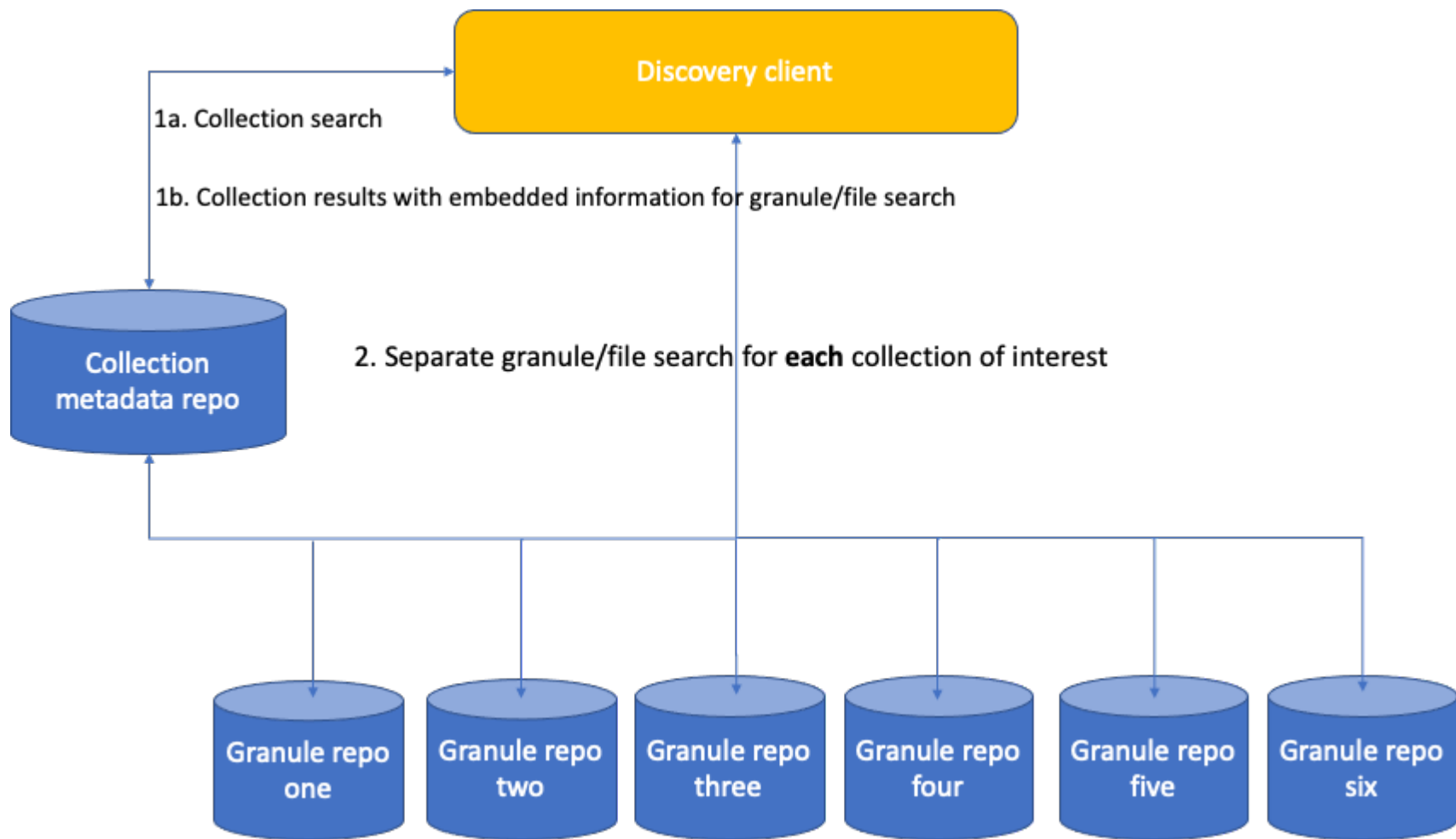
The CEOS WGISS Integrated Catalog (CWIC) provides the world's satellite Earth observation data providers with a way to make their data collections searchable using common standards.

CWIC architecture:

- consolidated collection metadata and discovery capability (CMR) **linked to...**
- distributed granule metadata and discovery and distribution capability

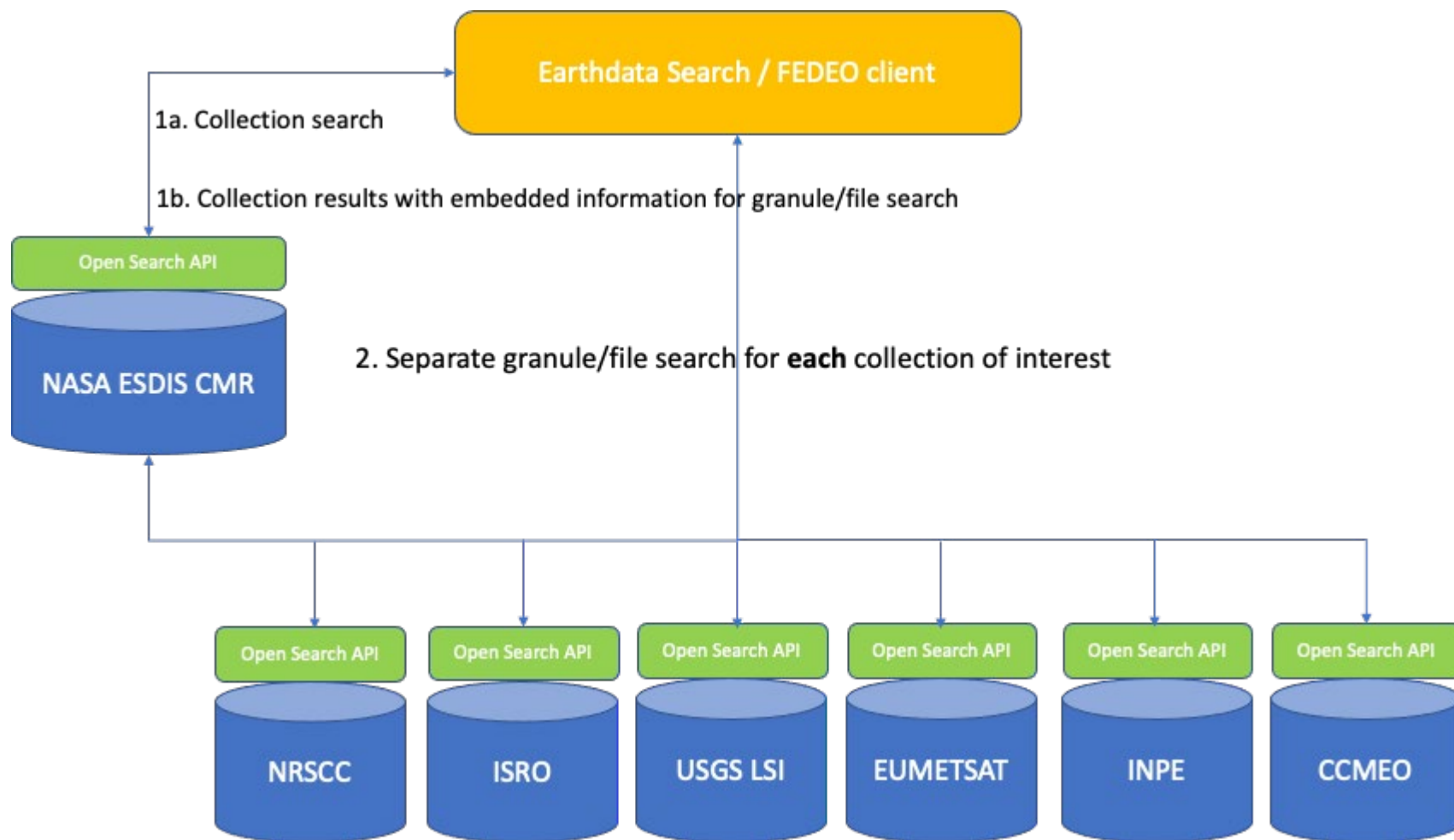
1. Find your collections of interest (CMR)
2. For each collection of interest
 - a. Find your granules of interest (agency API)
 - b. Access the granules (agency API)

Federated two-step searching





Open Search (not that one)



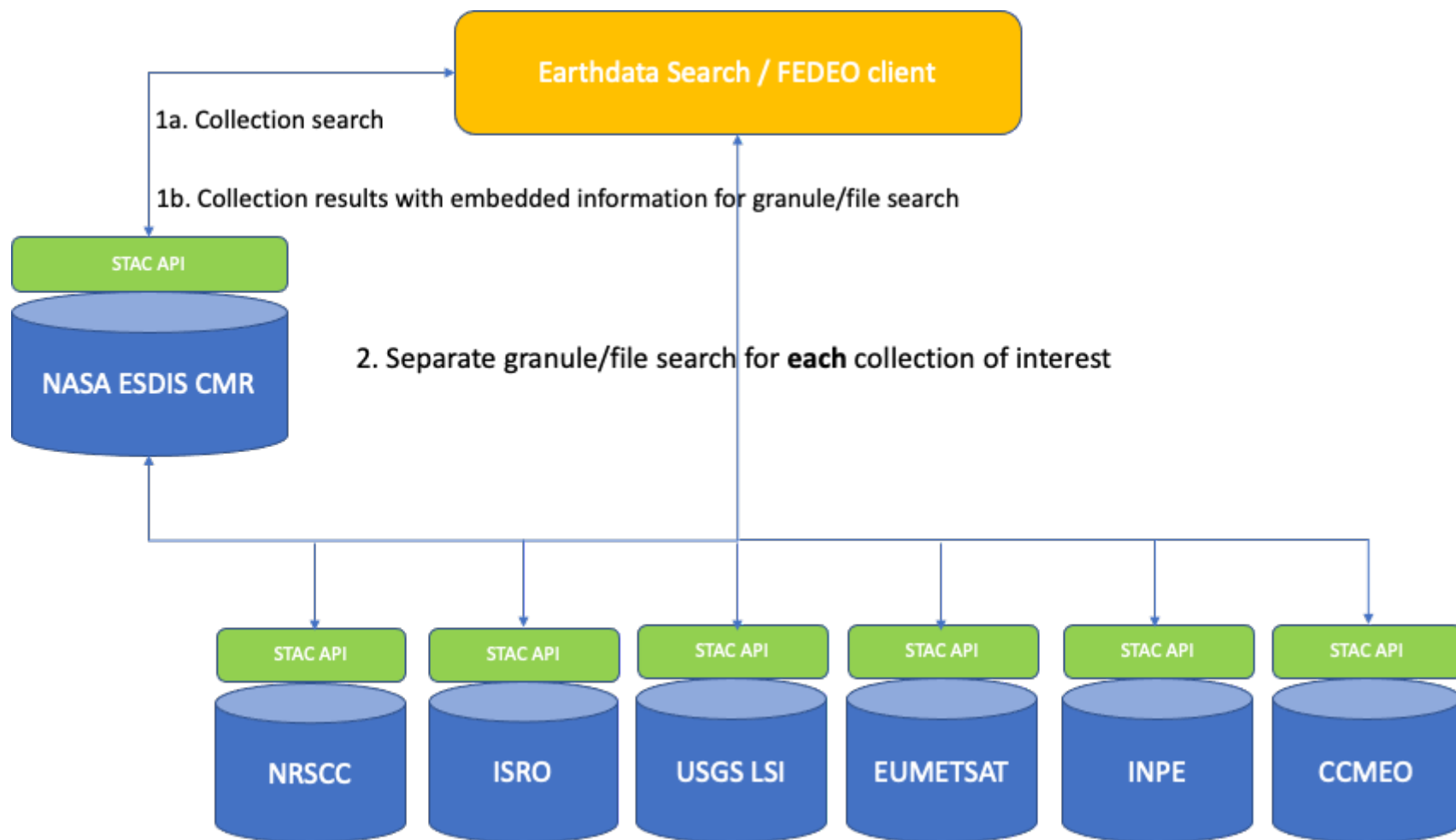
- Open Search had one solid use case
- Adoption was slow
- Operations and maintenance of agency APIs was of low quality
- Final straw was INPE deprecating their OpenSearch API to replace it with STAC
- STAC you say?



Enter STAC!

- ✓ Widely accepted standard
- ✓ Hypermedia As The Engine Of Application State (HATEOAS)
- ✓ 'collection' and 'items' implements two-step
- ✓ JSON result format
- ✓ Machine-readable API

```
{  
  "stac_version": "0.9.0",  
  "stac_extensions": [  
    "https://stac-extensions.github.io/eo/v1.0.0/schema.json",  
    "https://stac-extensions.github.io/projection/v1.0.0/schema.json"  
  ],  
  ...  
}
```



<https://cmr.earthdata.nasa.gov/stac/ISRO/collections>

collections:

```
[  
{
```

Standards based, JSON result format

Search step 1: collection-level search at NASA CMR

```
  id: "IMS1_HYSI_GEO.v1.0",  
  stac_version: "1.0.0-beta.2",  
  license: "not-provided",  
  title: "IMS-1 HYSI TOA Radiance and Reflectance Product",  
  description: "The data received from IMS1, HySI which operates in 64 spectral bands  
in VNIR bands(400-900nm) with 500 meter spatial resolution and swath of 128 kms.",  
  links:
```

HATEOAS - links including 'searching within this collection'

```
[  
  ...  
  {
```

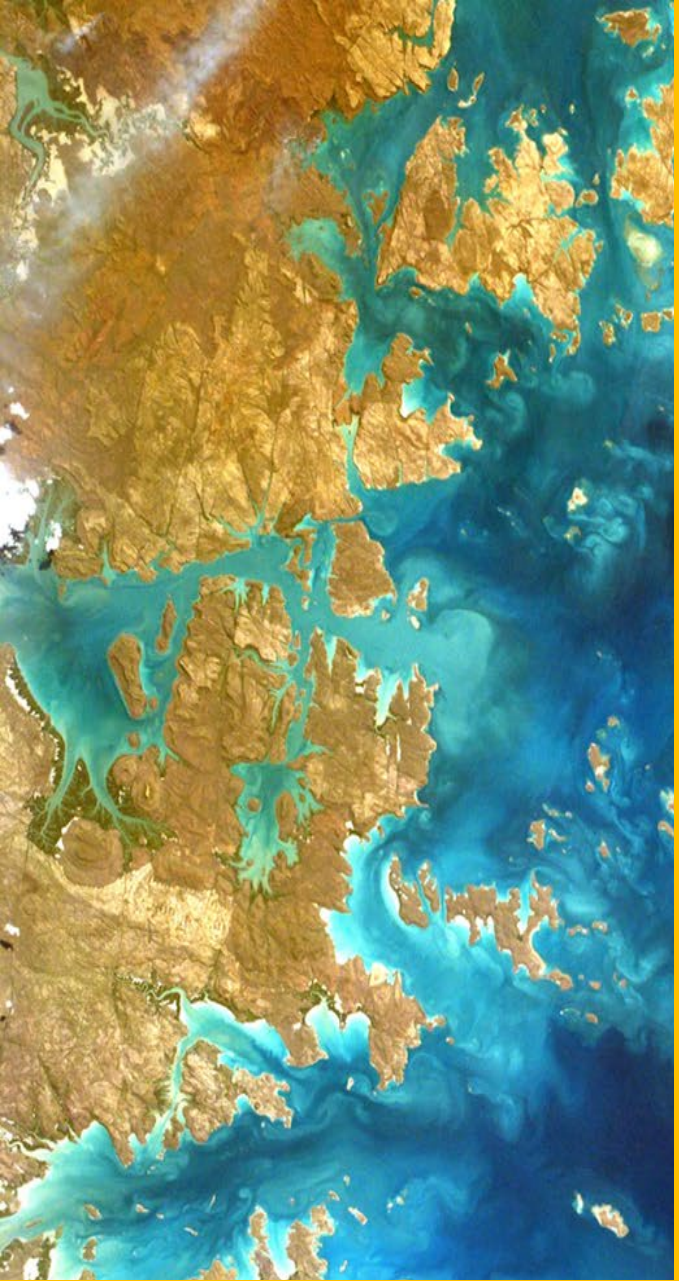
```
    rel: "items",  
    href:  
      "https://uops.nrsc.gov.in/stac/collections/IMS1_HYSI_GEO.v1.0/items",  
    title: "Granules in this collection",  
    type: "application/json"
```

Search step 2: file-level search at partner STAC API

Requires agency coordination to populate

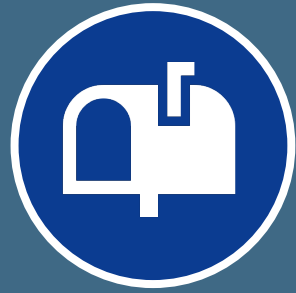
Machine-readable API from STAC standard

- Requires some [best practices](#) where choices are given
- ‘Two-step’ runs contrary to granule querying across collections
- Requires support of Collection Search extension
- Smart client needed to carry context from collection search to granule search (STACBrowser)



May the force be with you

Always



Email:
douglas.j.newman@nasa.gov



Website:
earthdata.nasa.gov



Social Media:
No idea