EOSDIS

NASA'S EARTH OBSERVING SYSTEM
DATA AND INFORMATION SYSTEM

STAC API for federated search

WGISS-54

October 4th, 2022

Doug Newman

NASA ESDIS Project, Goddard Space Flight Center

douglas.j.newman@nasa.gov

Summary

- A recap of the idea presented in WGISS-53
- Our candidate STAC extension
- Collection search
- Possible next steps
- Discussion

Open Search and STAC overview

Recap

Federate STAC in a nutshell

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

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

Standards based, **JSON** result format

Search step 1: collection-level search at NASA CMR

HATEOAS - links including 'searching within this collection'

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

Machine-readable API from STAC standard

EOSDIS

The case for STAC API

- Already solves use cases outside of the federated use case (cloud-friendly discovery, for example)
- Has all the features that made Open Search attractive for the implementation of a federated search architecture
- Some CWIC providers have developed their own STAC APIs (NASA, INPE)
- Existing, vibrant community and tooling (all open source)
  - STAC Browser
  - STAC API
  - pystac and intake-stac
- CMR and CMR STAC are open sourced and being used successfully by MAAP and NOAA NCEI
- STAC is specific to the earth science data domain

Needs for federated search

- A collection search endpoint definition
- Collection search parameter - free text keyword search
- An item search endpoint definition
- Providing programmatic linkage between distinct catalogs
  - Collections catalog -> Collection Items search

Federated search STAC extension

Extending the API

The proposed API extension

`https://github.com/littleidiot40/stac-federated-search`

The key points are as follows,

Federated linkage

Allowing links to Item Search endpoints from extra-domain catalogs in a catalog.

From CMR collections catalog:


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

Collection search

STAC API currently assumes the user knows the collection(s) the user requires and concentrates on item-specific dimensions.

To be a federated solution, we need to open up search at the collection level.

STAC and OGC collection searching

Decoupling collection search

STAC and collection search

OGC has a proposed API containing collection-level search support that suits our needs.

STAC is considering using that to solve the same problem.

Should we remove that from our extension and wait for / assume that STAC will implement something similar?

The OGC records API*

'The atomic unit of information in a catalogue is the record.

*A **catalog**ue is a **collection** of records.*
...

*A record provides a summary description of a resource that the provider of the resource wishes to make **discoverable**...'*

We wish to use 'parameter q' from the records API, a free text keywords parameter used to filter collections.

*http://docs.ogc.org/DRAFTS/20-004.htm

EOSDIS

STAC and OGC collection searching

Next steps

Get eyes on candidate extension

Initial review from Matt Hanson from STAC. Issues to be worked off

- Split out collection search
- Search endpoint definition
- Sundry corrections

We need more reviewers!

`https://github.com/littleidiot40/stac-federated-search`

Implement the federated extension

Once review is complete,

- Implement in CMR?
- Implement with Pathfinder CWIC provider (volunteers please!)

Stick or twist?

Do we wait for STAC to integrate with OGC wrt collection search?

Or do we forge ahead. If you build it, they will come!

Discussion