C-SCALE

Copernicus – eoSC AnaLytics Engine

EO Metadata Query Service

Efficient Discovery of Copernicus Data Assets

Stefan Reimond, EODC | Zdeněk Šustr, CESNET
stefan.reimond@eodc.eu | *sustr4@cesnet.cz*

C-SCALE | 22 March 2023

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017529.

Service Introduction

EO-MQS (Earth Observation Metadata Query Service)

- Evolved from the original plan to federate Copernicus Data providers within C-SCALE
  - **Discovery** across the federation a major goal
  - **Avoid** creating yet another matadata catalogue!
  - Data **access** is outside the scope of the EO-MQS
- Main premise: partners already know where their data are
  - Bring their discovery interfaces under a common one
    - single point
    - shared protocol
  - describe their datasets and data retention policies
  - use that to pre-select candidates and redistribute user queries

Datasets and Retention Policies

- Acknowledge that different partners have different data
  - National archives (full history, limited area)
  - Discipline archives (limited selection of product types, varying retention time)
  - Redistribution services (global coverage, short retention)
  - Big players (ambition to build global archive)

$\Longrightarrow$ Not every query needs to be redistributed to every partner

- Understand the query, select matching providers
- Currently taking into account only product type
  - *Area* and *time* filters not yet applied

Providers' Catalogue

- Called the *EO Resource Catalogue* initially $\rightarrow$ confusion
- Has **only** provider information, not data (product) information
  - Partners
  - Contacts
  - Services
  - Endpoints
- Adopting the well known GOC-DB (`https://goc.egi.eu/`)
  - The "Grid Configuration Database", put to new uses
  - Keep track of members and relevant service endpoints

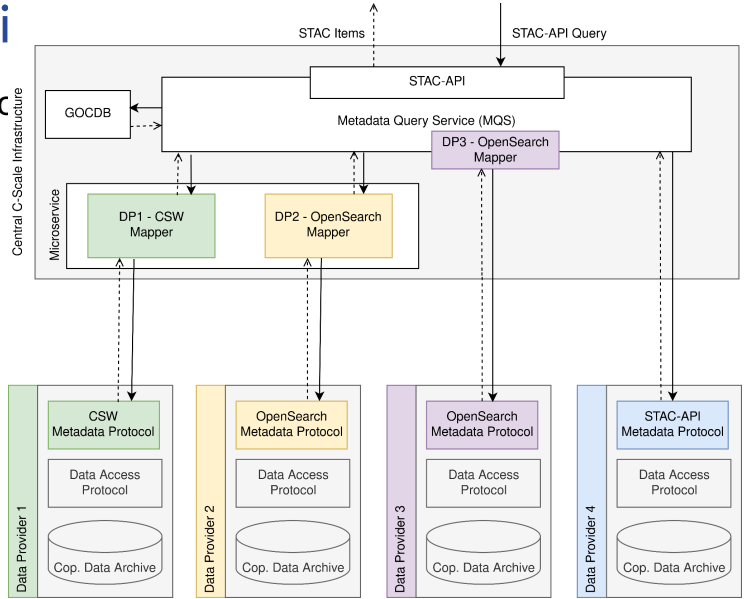
    `https://goc.egi.eu/gocdbpi/public/?method=get_service_endpoint&scope=C-SCALE`
  - Originally also intended for datasets and retention policies, but the choice of protocol changed that

Choosing the Common Protocol

- Choose from those already used in the federation?
  - OpenSearch, OData, CWS, STAC
  - Then implement translation
- **STAC** selected
  - modern protocol
  - lots of products supporting it
  - active community
  - `https://mqs.eodc.eu/stac/v1/`
- Features
  - Asset-focused: STAC breaks datasets into assets for easy data discovery.
  - Interoperability: STAC promotes compatibility with different tools and services.
  - Extensibility: STAC allows for custom metadata fields and extensions.

Choosi

- Choc
  -
  -


STAC Items
STAC-API Query
STAC-API
GOCDB
Metadata Query Service (MQS)
DP3 - OpenSearch Mapper
Central C-Scale Infrastructure
Microservice
DP1 - CSW Mapper
DP2 - OpenSearch Mapper
CSW Metadata Protocol
OpenSearch Metadata Protocol
OpenSearch Metadata Protocol
STAC-API Metadata Protocol
Data Access Protocol
Data Access Protocol
Data Access Protocol
Data Access Protocol
Cop. Data Archive
Cop. Data Archive
Cop. Data Archive
Cop. Data Archive
Data Provider 1
Data Provider 2
Data Provider 3
Data Provider 4

Demo: EO-MQS in Action

- Browsing and visualizing
  - Get to know the EO-MQS in your browser with STAC Browser and other tools
- Working with the EO-MQS in Python
  - Interact with EO-MQS programmatically in Jupyter notebooks.
- Extra: how to become a data prvoider?
  - Glance at the GOCDB

Let's get started $\rightarrow$ `https://mqs.eodc.eu/help`

Challenges

- Standardized STAC Collections structure?
  - At least for members who are building new STAC databases, it might make sense to use a common collection structure
- Paging
  - How to handle item paging when multiple backends respond?
  - Cache and collate own pages? Send more than the query asked for?
- Support of STAC API extensions?
  - The EO-MQS officially supports the STAC API Core parmaters
  - API Extension with filtering functionality is valuable, but hard to realize across federation

Outlook

- Constant improvement of EO-MQS package
  - New features
  - Fixing bugs
  - Better documentation
- Release of STAC Browser v3
  - Recently published, allows searching and filtering in the browser
  - Will be deployed soon at `https://mqs.eodc.eu/browser`
- EOSC Service Onboarding
  - The EO-MQS is presently undergoing onboarding as an EOSC service

Thank you

Questions?

Stefan Reimond, EODC | Zdeněk Šustr, CESNET
stefan.reimond@eodc.eu | *sustr4@cesnet.cz*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017529.