EUMETSAT – Effective Search and Access APIs for EO Data to Knowledge Value Chains

Uwe Voges (1), Michael Schick (2), Udo Einspanier (1), Gaetano Pace (3)

(1) https://www.conterra.de, u.(voges|einspanier)@conterra.de
(2) https://www.eumetsat.int, m.schick@eumetsat.int
(3) https://www.cgi.com/italy/en, gaetano.pace@cgi.com

Abstract. The EUMETSAT Data Services Roadmap included the implementation of an EO On-Line Data Access (OLDA) service offering search, discovery, visualization and access of EO products. conterra/N52 (part of a consortium led by CGI) have developed OpenSearch-EO- and REST-based Web API’s to browse, search and discover the products which are stored in a cloud-based storage system. The data is indexed for searching based on extended ISO/OGC metadata.

1. Introduction
The EUMETSAT Data Services Roadmap included the development of an Earth Observation (EO) On-Line Data Access (OLDA [1]) service allowing users and machine clients to search, visualize and download EUMETSAT EO products on the Web. This paper elaborates on OpenSearch-EO- and REST-based Web API’s conterra/N52 (part of a consortium led by CGI) have implemented in this context.

2. Organization, storage and indexing of EUMETSAT data and metadata in OLDA
EO collections in OLDA are sets of products, sharing a common collection metadata (e.g. platform or timeliness). Product metadata includes parameters like spatial/temporal extent or filename and is defined as profile of the OGC standard 10-157r4 [2]. A product is linked to a collection. Collection metadata is defined as profile of ISO 19115-2 and stored and indexed in ElasticSearch (https://www.elastic.co/) (via the Import/Export (IE) service). Products are ingested in the form of EUMETSAT Submission Information Packages (SIP). A SIP file includes a manifest file with structural information, the data object(s) (EO product, quicklook etc) and the EO product metadata. The SIP files are stored via the Storage Orchestrator (Fig.1) in a (federated) cloud-based storage system using object storage technology [3]. Upon SIP file ingestion, the EO product metadata is stored and indexed as well in ElasticSearch. To support grid-based spatial filters, the metadata is indexed based on the Equi7 Grid tiling system [3].

3. Browse, search, discovery and access Web API’s in OLDA
To easily browse, search, discover and access the EO collections and products three modern Web API’s are provided (used by OLDA UI (Fig.1)): REST-EO API, OpenSearch-EO API, Download API (not further considered here, see [3]). These APIs are implemented as microservices (docker container) to guarantee a high degree of automation. Scalability and reliability are overtaken by Kubernetes [3]. WSO2 API Gateway serves as single-entry point into OLDA and is responsible for authorization (products have different access restrictions). Authentication and UserManagement services are based on WSO2 Identity Manager, which enables clients to access the APIs using the standard OAuth2 protocol.

The REST-EO API is a simple to use REST-API to browse products by humans (HTML), search engines (HTML+schem.org) or other machine clients (GeoJSON). Within a data to knowledge
value chain (e.g. hosted processing using JupyterNotebook) an API offering simple iterations through “virtual directories” providing huge amounts of products is more beneficial than a search interface.

![Figure 1. EUMETSAT Online Data Access (OLDA) – High Level Architecture, UI snippets](image)

At its root path the API provides the available collections. Starting from a collection a client can discover collection metadata or browse through the assigned products based on a predefined set of axes (temporal, spatial). During this navigation, the set of remaining products is narrowed step by step by drilling into hierarchically organized subsets. The results of a subset are not thousands of products but rather the next finer grained subsets. Just for subsets with a manageable number of remaining products links to the products are provided (e.g. 2018 -> 01-2018 -> 08-01-2018 -> products). By following these links, metadata details and data access (download) links are provided. The API is aligned with OGC/W3C Spatial Data on the Web Best Practices ([https://www.w3.org/TR/sdw-bp](https://www.w3.org/TR/sdw-bp)) and OGC API Common ([www.opengeospatial.org](http://www.opengeospatial.org)) and follows principles from well-known API’s like Amazon’s S3.

For flexible searches along varying axes OpenSearch-EO [4] is implemented. OpenSearch is a standards-based web search interface providing an interface description, the OpenSearch Description Document (OSDD). The OSDD includes one URL template per supported response format (e.g. JSON) with key-value-pairs (defining the type and value range) to constrain the search. On the one hand OpenSearch-EO extends OpenSearch by EO-related query parameters. For collection search in OLDA we only support full-text search. For product search it’s possible to filter on the mission, TOI and AOI. A search result includes the search entries (metadata about the original item) and information about the current search. OpenSearch-EO extends the response by EO related elements and a GeoJSON encoding (both implemented in OLDA). The entries include links to external information (in OLDA e.g. to GeoJSON encoded product metadata (as defined in OGC 17-003 [4]) and to the data access).

References
[1] EUMETSAT On-Line Data Access (OLDA): [https://olda.eumetsat.int/olda-ui/](https://olda.eumetsat.int/olda-ui/)
[2] OGC Earth Observation Metadata profile of Observations & MeasurementsStandard (OGC 10-157r4. [https://docs.opengeospatial.org/is/10-157r4/10-157r4.html](https://docs.opengeospatial.org/is/10-157r4/10-157r4.html))
[3] EO ON-LINE DATA ACCESS IN THE BIG DATA ERA (2019): EO On-Line Data Access (OLDA) in the Big Data era. Joint Research Center (EU Science Hub), [http://publications.jrc.ec.europa.eu/repository/bitstream/JRC115761/procbids19.pdf](http://publications.jrc.ec.europa.eu/repository/bitstream/JRC115761/procbids19.pdf)
[4] OGC OpenSearch-EO 1.1 (OGC 13-026r9, 17-047. 17-003). [www.opengeospatial.org](http://www.opengeospatial.org)