www.eumetsat.int

M. Ziółkowski[1], J. Musiał[1], J. Bojanowski[1], P. Mujta[1], M. Bylicki[1], A. Lambare[2]
A. Le Carvennec[2], T. Hilton[2], C. Reimer[3], M. Schick[4]

DestinE Data Lake – Harmonized Data Access (HDA)

INTRODUCTION

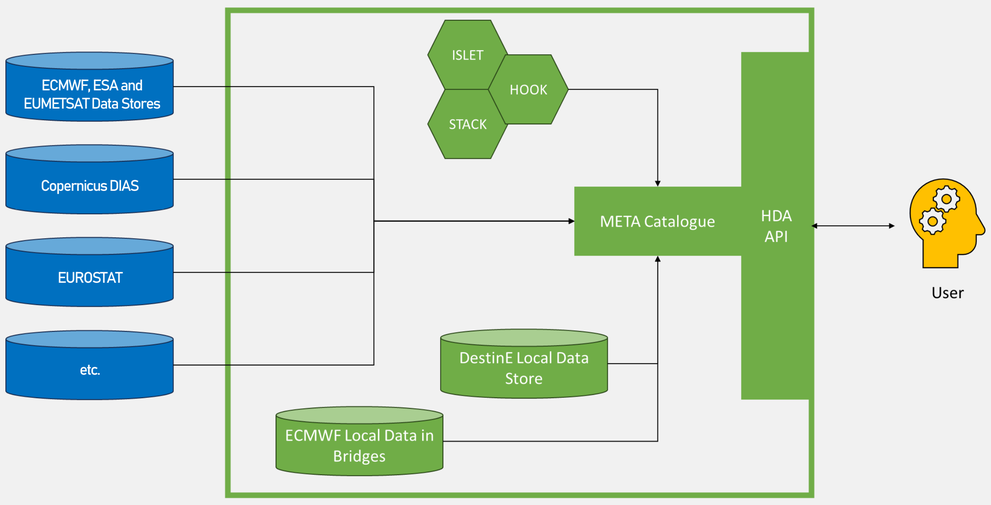
The DestinE Data Lake's HDA is the API to discover and access all data available in the DestinE Data Portfolio. The HDA is based on the federation principle make use of Federated Data Provider data offerings as well as colocalised data within the DestinE Data Lake and DEDL services.

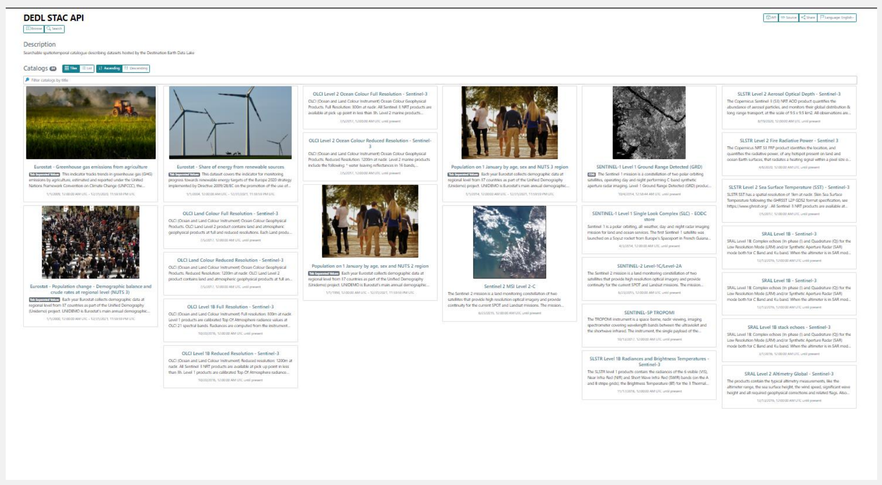
Federation of Data Sources

The architecture of the Data Lake HDA is exposing a STAC 1.0.0 compliant interface based on a meta catalogue (instance of EOCATALOGUE) describing the collections and relying on EODAG to access the federated data sources. The benefits are:

- No duplication of data
- No systematic harvesting of data or meta data of the products
- Multiple data provider sources can be referenced for a given collection improving the robustness of the solution

The HDA provides a consistent STAC interface to discover and access the data whatever the federated data provider source protocol and methodology of data access. The code used to access and manipulate the data becomes independent of the data source. The HDA also enables do discover and access other DEDL services.



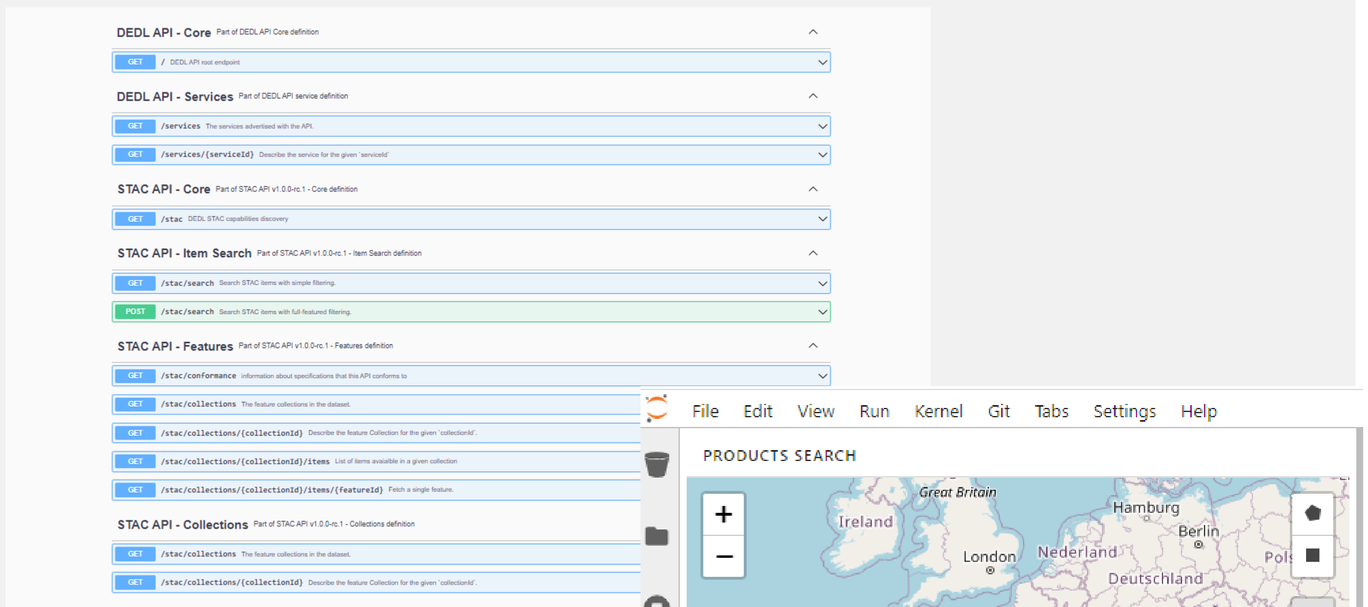


Access with: STAC API

The standard access of the HDA uses the STAC interface to discover and access the data with additional functionalities as synchronous and asynchronous data download. PyStac client can be used.

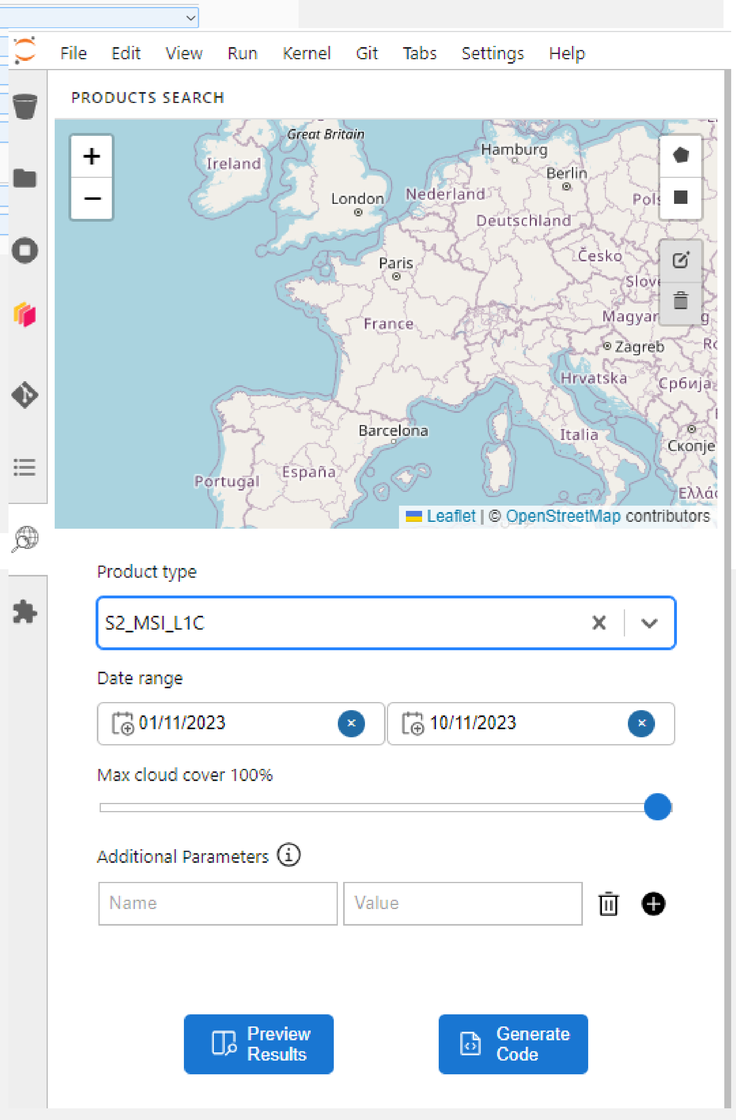
```
# AOI around Delfzijl, in northern Netherlands
# limit sets the # of items per page so we can see multiple pages getting fetched
search = cat.search(
    max_items = 20,
    limit = 10,
    collections = "EO.ESA.DAT.SENTINEL-1.L1_GRD",
    bbox = [-11,35,50,72 ],
    datetime = "2023-09-09/2023-09-20",
)

for item in search.items():
    print(item.id)
```

```
S1A_EW_GRDM_1SDH_20230912T034103_20230912T034203_050289_060DD6_F440.SAFE
S1A_EW_GRDM_1SDH_20230912T034203_20230912T034303_050289_060DD6_C034.SAFE
```



Access with: PYTHON Library

The EODAG library & JupyterLab extension is available from DEDL services to make it even more simple for accessing Data from Python.



Leveraging existing tools

The HDA provides a STAC interface and any tools supporting STAC can be configured to discover, visualize and manipulate the data available in the DestinE Data Lake. For example STAC BROWSER, QGIS with STAC extension…

Open Source

EODAG and EOCATALOGUE are Apache 2.0 Open source Software

CloudFerro
a Sopra Steria company
eodc
EUMETSAT
Funded by the European Union
Destination Earth
IMPLEMENTED BY EUMETSAT esa ECMWF

References

[1] CloudFerro S.A. Nowogrodzka 31, Warsaw Poland
[2] CS Group, avenue Galilée, LE Plessis Robinson, France
[3] EODC, Franz-Grill-Straße 9, Vienna, Austria
[4] EUMETSAT, Eumetsat-Allee 1, Darmstadt, Germany
[5] EODAG - https://github.com/CS-SI/eodag
[6] EOCATALOGUE -
[7] STAC - https://stacspec.org/