



# Harmonized Data Access service

A unified API for DestinE Data Portfolio

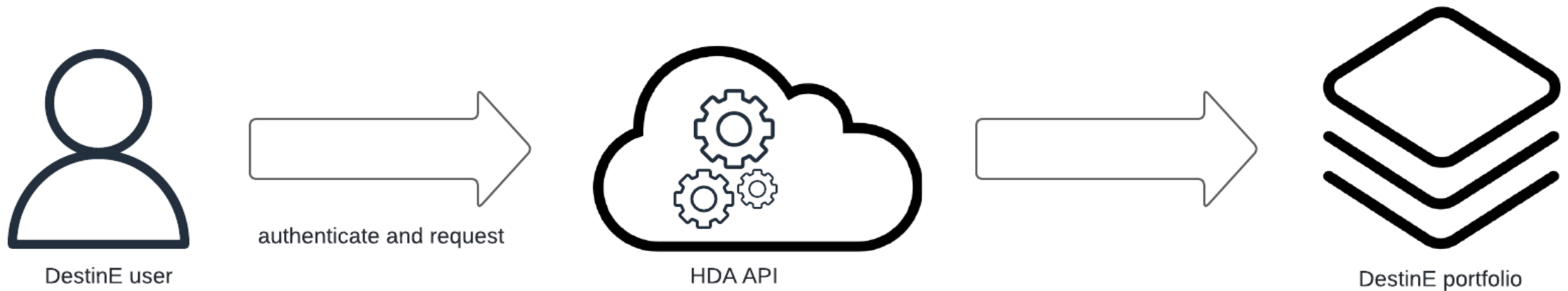


# A Harmonized Data Access

- **1** set of credentials

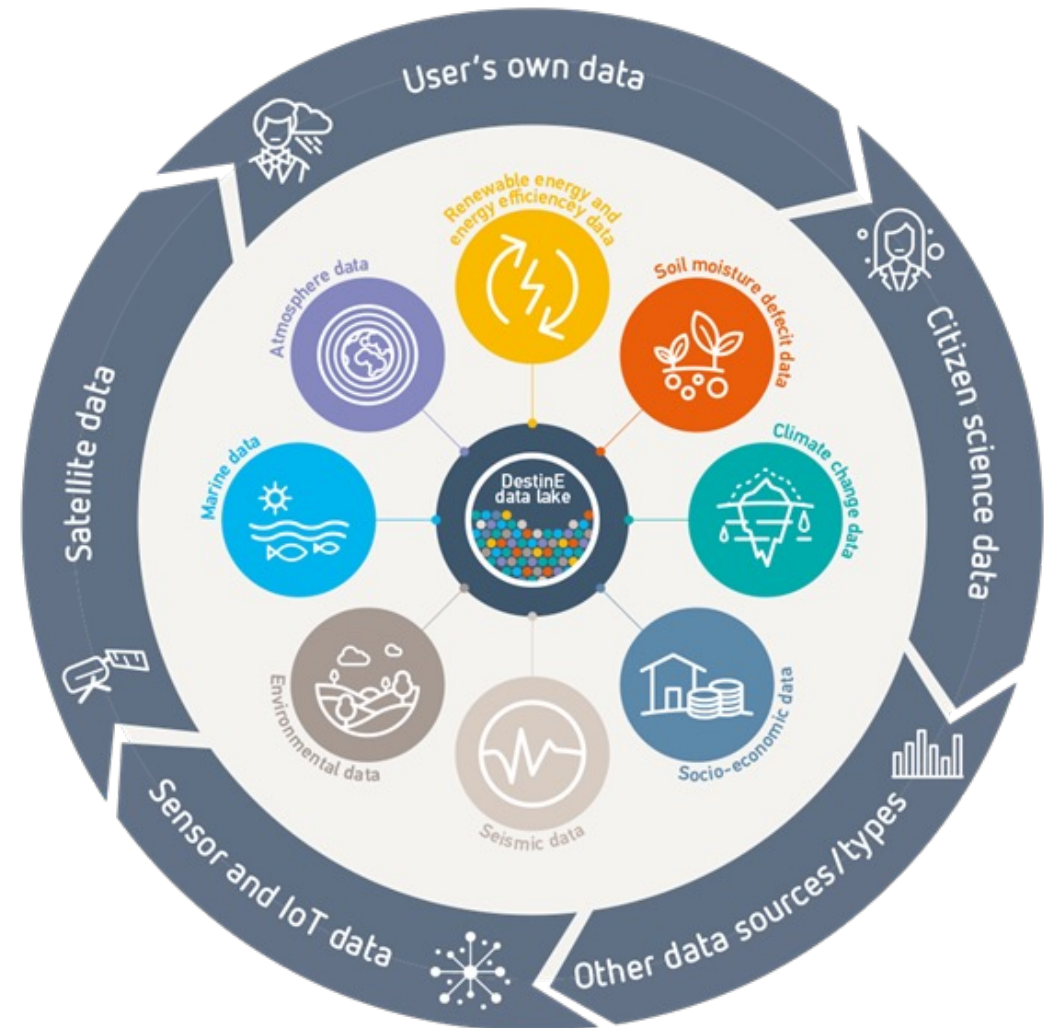
- **1** API

- More than **150** datasets



# A diverse data portfolio

- Earth Observation (EO)
- Weather, Climate, Hydrology
- Survey, statistics and knowledge
  - Ground
  - Population
  - Biological
  - Socio-economic data
- Digital Twin Outputs



# A federation of data providers

- Federation first
- A local cache
- Redundancy



# A standard interface

- SpatioTemporal Asset Catalog (STAC)
- Filter by
  - Temporal extent
  - Spatial extent
  - Sort
  - Custom properties

## STAC API - Core Part of STAC API v1.0.0-rc.3 - Core definition

GET `/stac` DEDL STAC capabilities discovery

GET `/stac/conformance` information about specifications that this API conforms to

## STAC API - Collections Part of STAC API v1.0.0-rc.3 - Collections definition

GET `/stac/collections` The feature collections in the dataset.

GET `/stac/collections/{collectionId}` Describe the feature Collection for the given `collectionId`.

## STAC API - Features Part of STAC API v1.0.0-rc.3 - Features definition

GET `/stac/collections/{collectionId}/items` List of items available in a given collection

GET `/stac/collections/{collectionId}/items/{featureId}` Fetch a single feature.

## STAC API - Item Search Part of STAC API v1.0.0-rc.3 - Item Search definition

GET `/stac/search` Search STAC items with simple filtering.

POST `/stac/search` Search STAC items with full-featured filtering.

# Discover datasets that matter

- Intuitive data discovery
- Filter by
  - Temporal extent
  - Title, description, keywords
  - Categories
- A web portal

<https://hda.data.destination-earth/ui>

The screenshot shows the DestinE Data Portfolio web portal. At the top, there is a navigation bar with the Destination Earth logo and links for Data Lake, DestinE Data Portfolio, Documentation, and Edge Services. A 'Sign In' button is located in the top right corner. Below the navigation bar, the main heading is 'DestinE Data Portfolio'. A sub-heading reads: 'The DestinE data lake federates with existing data holdings as well as with complementary data from diverse sources like in-situ, socio-economic, or data-space data.' Below this is a search bar labeled 'Filter datasets'. The main content area is titled 'Air Quality' and features a vertical sidebar on the left with a list of categories: Air Quality, Atmosphere, Biomass/Vegetation, Climate, Climate/Weather, Climate/Wind, DEMs, Demographic, Fire, Ice, Imagery, Land cover, Other, Soil, Solar, Temperature, Water, and Wind. The main content area displays several dataset cards, each with a map thumbnail, a title, a description, and a set of filter buttons. The cards shown are: 1. 'CAMS European air quality forecasts' with a map of Europe and filters for Aerosol, Europe, airPollution, Atmospheric conditions, Reactive gas, and Past. 2. 'CAMS European air quality reanalyses' with a map of Europe and filters for Aerosol, Reanalysis, Europe, airPollution, Atmospheric conditions, and Reactive gas. 3. 'CAMS global atmospheric composition forecasts' with a world map and filters for Aerosol, Global, airPollution, Atmospheric conditions, Reactive gas, and Past. 4. 'CAMS global biomass burning emissions based on fire radiative power (GFAS)' with a world map and filters for Aerosol, Global, airPollution, Atmospheric conditions, Reactive gas, and Past. 5. 'CAMS global emission inventories' with a world map and filters for Aerosol, Global, airPollution, Atmospheric conditions, Reactive gas, and Past. 6. 'CAMS global greenhouse gas reanalysis (EGG4)' with a world map and filters for Global, Reanalysis, airPollution, Atmospheric conditions, and Past.

# Identify available filters

- Queryables API

**GET** `/stac/queryables` Get the JSON Schema defining the list of variable terms that can be used in CO

**GET** `/stac/collections/{collectionId}/queryables` Get the JSON Schema defining the list

Chosen collection: EO.ECMWF.DAT.CAMS\_SOLAR\_RADIATION\_TIMESERIES

Fetch queryables

sky\_type

time\_refere...

format

Properties fetched successfully.

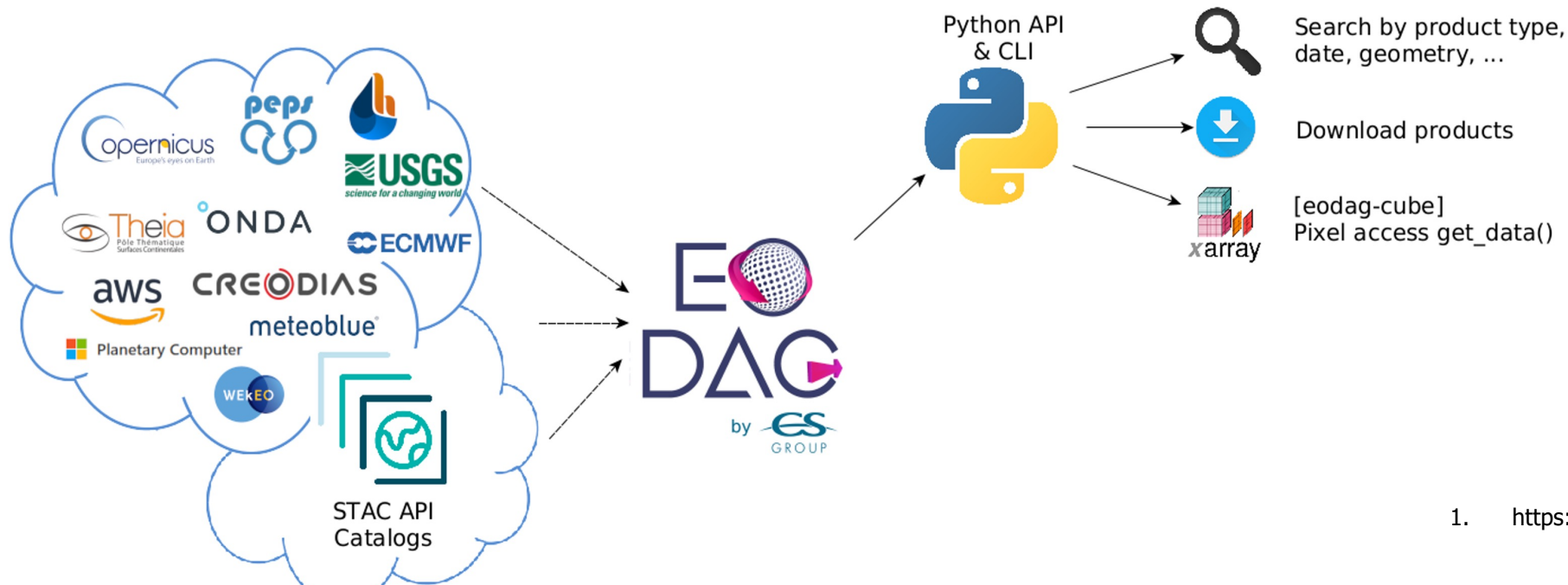
Applicable filters

Description	Type	enum	value
sky_type	"string"	clear , observed_cloud	"clear"
time_reference	"string"	true_solar_time , universal_time	"true_solar_time"
format	"string"	csv , netcdf	"csv"

```
{
  "$schema": "https://json-schema.org/draft/2019-09/schema",
  "$id": "http://localhost:8088/collections/DT_CLIMATE_ADAPTATION/queryables",
  "type": "object",
  "title": "Queryables for DestinE Data Lake STAC API",
  "description": "Queryable names for the DestinE Data Lake STAC API Item Search",
  "properties": {
  },
  "ecmf:type": {
    "const": "fc",
    "default": "fc",
    "title": "type"
  },
  "ecmf:model": {
    "const": "IFS-NEMO",
    "default": "IFS-NEMO",
    "title": "model"
  },
  "ecmf:resolution": {
    "default": "high",
    "enum": [
      "high",
      "standard"
    ],
    "title": "resolution",
    "type": "string"
  },
  "ecmf:param": {
    "default": "134/165/166",
    "enum": [
      "134",
      "137",
      "141",
      "144",
      "146",
      "147",
      "148",
      "151",
      "159",
      "164",
      "165",
      "166",
      "167",
      "168",
    ]
  }
}
```

# Privileged access with EODAG<sup>1</sup>

- Command line tool
- Plugin-oriented Python framework
- Available on Github (Apache 2.0)
- Extensions (eodag cube, JupyterLab...)
- Each data provider is a plugin
  - Ease to extend with tutorials and documentation



1. <https://github.com/CS-SI/eodag>



# Privileged access with EODAG – Usage examples

- **Python**

```
from eodag import EODataAccessGateway

dag = EODataAccessGateway()

search_results, total_count = dag.search(
    productType="S2_MSI_L1C",
    start="2021-03-01", end="2021-03-31",
    geom={"lonmin": 1, "latmin": 43, "lonmax": 2, "latmax": 44}
)
product_paths = dag.download_all(search_results)
```



- **CLI**

```
eodag search --productType S2_MSI_L1C --box 1 43 2 44 --start 2021-03-01 --end 2021-03-31
eodag download --search-results search_results.geojson
```

# Privileged access with EODAG - labextension



**PRODUCTS SEARCH**

Product type

Date range

Max cloud cover 75%

Additional Parameters ⓘ

**Search results**

**Quicklook**

**Metadata**

Platform serial identifier	S2B
Instrument	MSI
Product type	S2MSI1C
Processing level	LEVEL1C
Sensor mode	INS-NOBS
Start time from ascending node	2023-08-08T10:46:29.024Z
Completion time from ascending node	2023-08-08T10:46:29.024Z

Start time	End time	Cloud cover	Actions
2023-08-08T10:46:29.024Z	2023-08-08T10:46:29.024Z	0 %	🔍
2023-08-06T10:56:21.024Z	2023-08-06T10:56:21.024Z	21 %	🔍
2023-08-05T10:36:29.024Z	2023-08-05T10:36:29.024Z	27 %	🔍
2023-08-03T10:46:31.024Z	2023-08-03T10:46:31.024Z	35 %	🔍

```
[1]: from eodag import EODataAccessGateway, setup_logging
      setup_logging(1) # 0: nothing, 1: only progress bars, 2: INFO, 3: DEBUG

      dag = EODataAccessGateway()
      geometry = "POLYGON ((1.054688 47.631835, -2.548828 43.576146, 0.439453 40.642858, 5.712891 44.083376, 1.054688 47.631835))"
      search_results, total_count = dag.search(
        productType="S2_MSI_L1C",
        geom=geometry,
        start="2023-08-01",
        end="2023-08-10",
        cloudCover=75,
        **{
          "tileIdentifier": "31TCJ",
        }
      )

[*]: path = dag.download(search_results[1])

S2A_MSIL1C_20230806T105621_N0509_R094_T31TCJ_20230806T144314: 38% 12.3M/32.0M [00:22-00:31, 624KB/s]
```

# Native integration with all STAC tools

- STAC browser<sup>1</sup>
- The python library PySTAC<sup>2</sup>
- QGIS STAC API Browser<sup>3</sup>

1. <https://github.com/radiantearth/stac-browser>
2. <https://github.com/stac-utils/pystac>
3. <https://github.com/stac-utils/qgis-stac-plugin>

The screenshot shows a web interface for a STAC catalog. At the top, there are navigation options like 'Catalogues 151', 'Toules', 'Liste', and sorting options 'Ordre croissant' and 'Ordre décroissant'. Below this is a search filter: 'Filtrer les catalogues par titre, description ou mots-clés'. The main area displays a grid of product cards, each with a thumbnail and a brief description. The cards include:

- AMSU-A Level 1B - Metop - Global**: The Advanced Microwave Sounding Unit-A (AMSU-A) is a 15-channel microwave radiometer that is used for measuring global atmospheric... (01/03/2008 00:00:00 UTC jusqu'à présent)
- CAMS global reanalysis (EAC4) monthly averaged fields**: EAC4 (ECMWF Atmospheric Composition Reanalysis 4) is the fourth generation ECMWF global reanalysis of atmospheric composition... (01/01/2003 00:00:00 UTC - 31/12/2022 23:59:59 UTC)
- Glacier Distribution**: Worldwide distribution of glacier outlines associated with individual glacier parameter including hypsometry. The inset figures show a close up of the outlines of the Rhone glacier in Switzerland and the corresponding hypsometry. (01/01/2000 00:00:00 UTC - 31/12/2000 23:59:00 UTC)
- ASCAT Coastal Winds at 12.5 km Swath Grid - Metop**: Equivalent neutral 10m winds over the global oceans, with specific sampling to provide as many observations as possible near the coasts. Better... (16/04/2013 00:00:00 UTC jusqu'à présent)
- CAMS solar radiation time-series**: The CAMS solar radiation services provide historical values (2004 to present) of global (GHI), direct (BHI) and diffuse (DHI) solar irradiation, as... (01/02/2004 00:00:00 UTC jusqu'à présent)
- Glaciers elevation and mass change data from 1850 to present from the Fluctuations of Glaciers Database**: This dataset provides in situ and remote sensing derived glacier changes from individual glaciers globally. The dataset represents the latest... (01/01/1850 00:00:00 UTC - 31/12/2019 00:00:00 UTC)
- Global Ocean Mean Dynamic Topography**: Mean Dynamic Topography that combines the global CNES-CLS-2022 MDT, the Black Sea CMEMS2020 MDT and the Med Sea CMEMS2020... (01/01/2003 00:00:00 UTC)
- Global Ocean Monthly Mean Sea Surface**
- MHS Level 1B - Metop**: The Microwave Humidity Sounder instrument used to provide retrieval of surface temperature... (23/03/2009 00:00:00 UTC jusqu'à présent)
- Multi Observation Global Temperature Salinity Heat Current and**: You can find here the Multi Observation ARMOR3D L4 analysis reprocessing. It consists of 3D... (01/01/1993 00:00:00 UTC - 05/01/2022 00:00:00 UTC)

<https://radiantearth.github.io/stac-browser/#/external/hda.data.destination-earth.eu/stac>

# Key takeaways

- Interface between users and the DestinE Data Lake portfolio
- Harmonized data access API
- Federation of providers
- STAC compliant
- EODAG library, CLI and JupyterLab extension

# Thank you!

## Q&A

---

aubin.lambare@cs-soprasteria.com

API



<https://hda.data.destination-earth.eu>

Python tutorials



<https://github.com/destination-earth/DestinE-DataLake-Lab>

Web portal



<https://hda.data.destination-earth.eu/ui>

Documentation



<https://destine-data-lake-docs.data.destination-earth.eu>



a Sopra Steria company