

DeepESDL – an introduction

Tejas Morbagal Harish, Brockmann Consult GmbH
19 February 2024

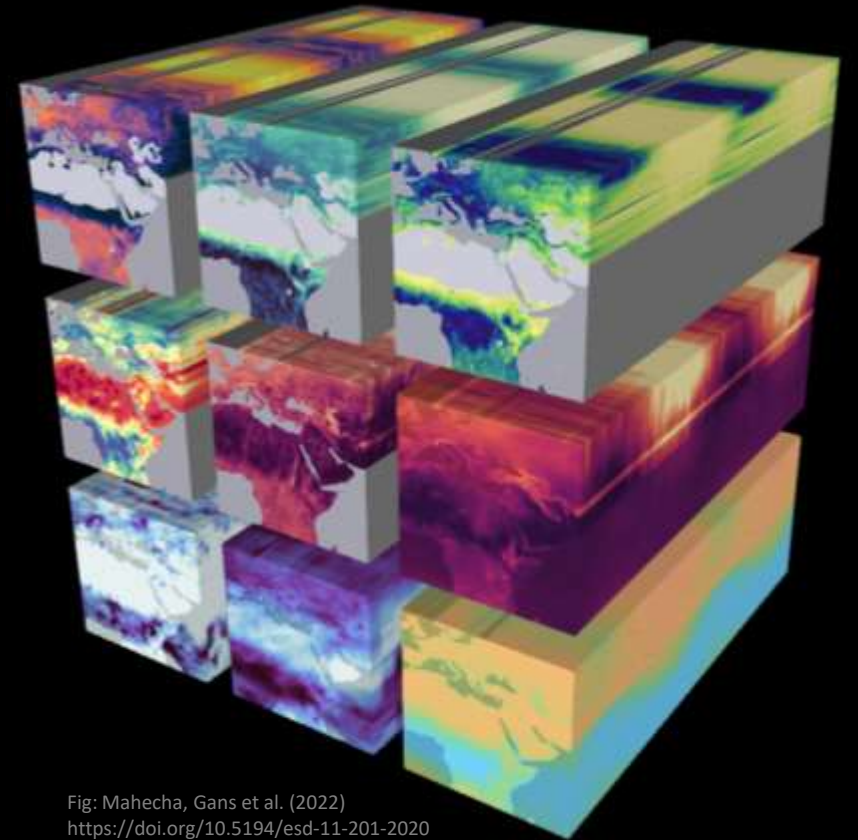


Fig: Mahecha, Gans et al. (2022)
<https://doi.org/10.5194/esd-11-201-2020>

What is DeepESDL?

*“DeepESDL” shall serve the needs of ESA **scientific exploitation** projects, the **Science Clusters** and in particular the **Science Hub**. In particular it shall allow different scientists and **teams** to develop and exploit tailored **multi-variate, multi-temporal, local to regional and global scale thematic datasets in different levels of scientific analysis**.*

- An **integrator of scientific results** from different activities into a single infrastructure.
- A platform for **collaborative** and **reproducible research** allowing different scientists and teams to work together in a collective scientific effort, sharing data, tools and expertise.
- Support for the **execution of individual projects** from ESA or from the scientific community worldwide.

DeepESDL Documentation

The screenshot shows the homepage of the DeepESDL documentation. The header includes a search bar and navigation links for HOME, About DeepESDL, USER GUIDE, DATASETS, and DESIGN. The main content area features a large image of a data cube and the text "EARTH SYSTEM DATA LAB". Below this is the "About DeepESDL" section, which provides a welcome message and a list of contents.

About DeepESDL

Welcome to the online documentation of DeepESDL - ESA's Deep Earth System Data Laboratory, a platform providing analysis-ready data cube in a powerful, virtual laboratory to the Earth Science research community. DeepESDL offers a full suite of services to facilitate data exploitation, share data and source code, and publish results. Special emphasize is put on improving the support for machine learning and artificial intelligence approaches, which includes the preparation of AI-ready datasets, providing a programming environment with relevant libraries and packages, and the resources to execute processing pipelines. For more information and access to the lab please visit the [DeepESDL website](#).

The DeepESDL documentation contains:

- The [User Guide](#) for all DeepESDL services. This is the starting point for new users.
- An overview of [public, pre-generated data cubes](#) available in DeepESDL with detailed metadata and specifications.
- A [description of DeepESDL's architecture](#).

Copyright © 2022 by the DeepESDL Team

This screenshot shows the "About DeepESDL" page, which is part of the "User guide" section. It includes a search bar, navigation links, and a table of contents. The main content area is titled "Earth System Data Cube (ESDC) v3.0.0" and includes a globe image and a table of parameters.

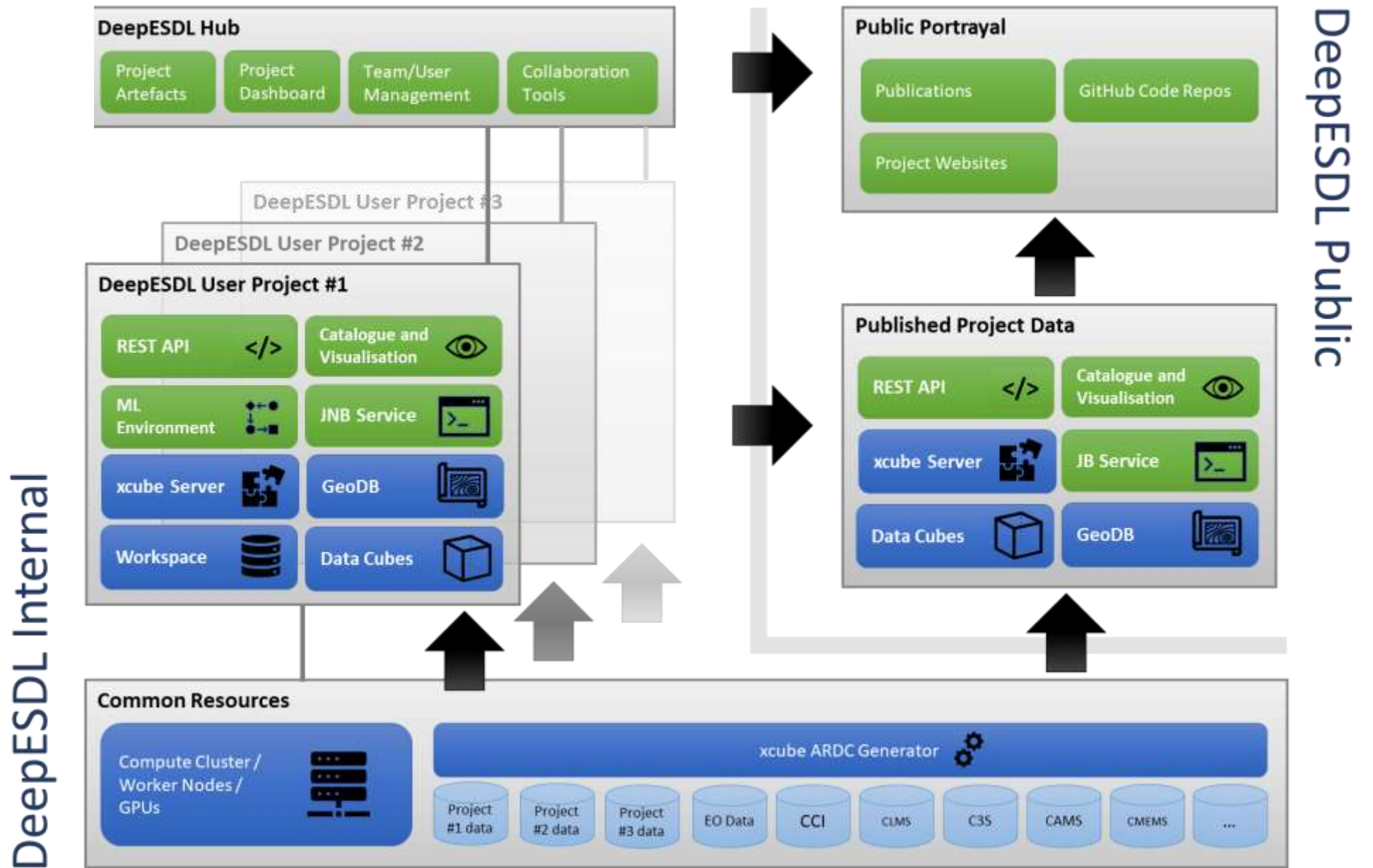
Earth System Data Cube (ESDC) v3.0.0

Basic information

Parameter	Value
Bounding box latitude	-89.875 to 89.875
Bounding box longitude	-179.875 to 179.875
Time range	1979-01-05T00:00:00.000000000 to 2021-12-31T00:00:00.000000000
Time period	3D
Publisher	DeepESDL Team

[Click here for full dataset metadata.](#)

Architecture Overview



Service Offer

- Configurable **Jupyter** Hub with user-defined **Python** environments based on the common data science stack
- **Reproducible workflows** can be implemented as Jupyter notebooks.
- **Compute** resources including GPU cores.
- Ever-growing data offer with relevant Earth Science data, as **persisted** Analysis-ready Data Cubes or as **virtual**, remote cubes via xcube data stores (CCI, CMEMS, C3S, Sentinel Hub, ...)

Now let's search for snow datasets provided via the xcube cci store:

```
[9]: iterator = cci_store.search_data(cci_attrs=dict(ecsv='SNOW'))
JSON([item.to_dict() for item in iterator])
```

[9]: **root:** [] 18 items

- ▶ **0:**
 - ▶ **attrs:**
 - ▶ **bbox:** [] 4 items
 - ▶ **coords:**
 - crs: "\WGS 84"
 - data_id: "esacci...
 - data_type: "datas...
 - ▶ **data_vars:**
 - ▶ **dims:**
 - ▶ **open_params_schem**
 - spatial_res: 0.81
 - time_period: "1D"
 - ▶ **time_range:** [] 2

DeepESDL Documentation

Earth System Data Cube (ESDC) v3.0.1

How to open this dataset in DeepESDL JupyterLab

```
from xcube.data_store import new_data_store
store = new_data_store('CCI', '1900-01-01-CCI-pub111', storage_opts=dict(aws=True))
ds = store.open_data('esacci-0-2024-12-31-01-01-01')
```

Bounding box map

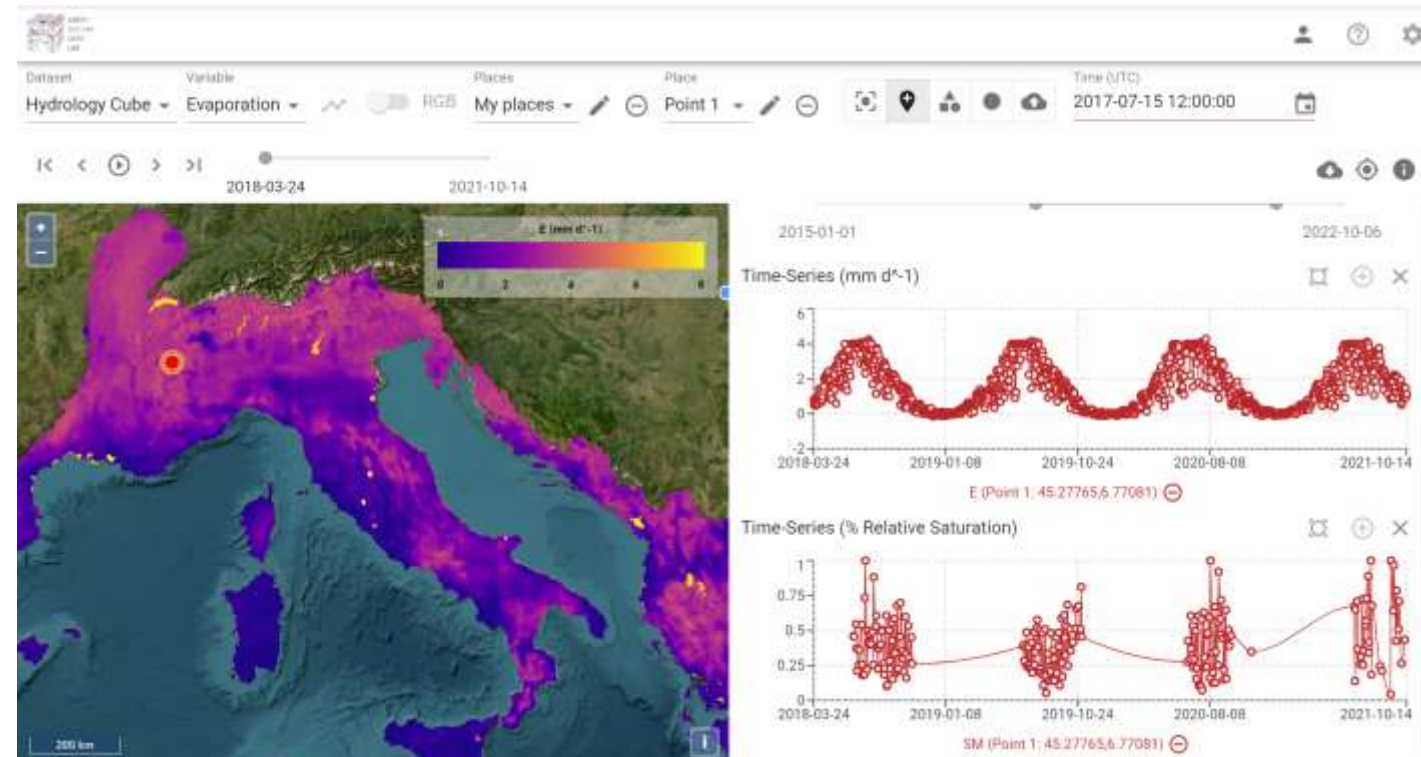
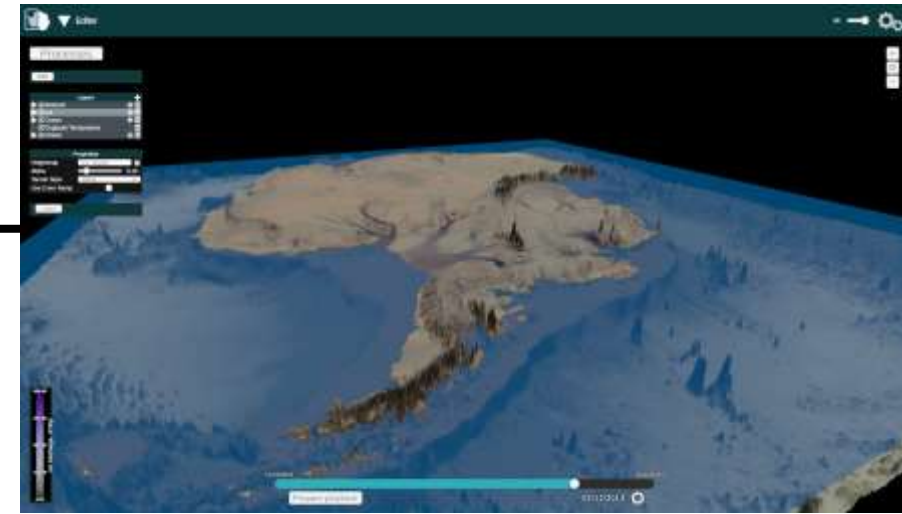
Basic information

Parameter	Value
Bounding box longitude (°)	-180.0 to 180.0
Bounding box latitude (°)	-90.0 to 90.0
Time range	1979-01-05 to 2023-12-31

<https://deepesdl.readthedocs.io/en/latest/guide/jupyterlab/#creating-custom-team-python-environment>

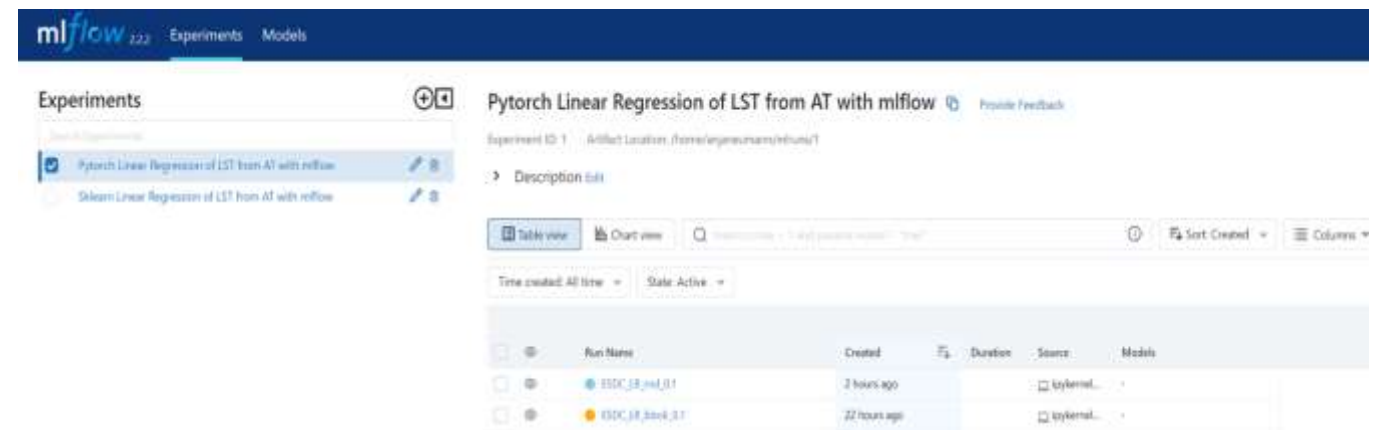
Service Offer

- Data **Catalogue** for DeepESDL Analysis-Ready Data Cubes
- Different **visualization** solutions for exploration and dissemination of results (xcube viewer, 4D viewer, Lexcube)
- Optional services upon request like geoDB and Sentinel Hub
- Scalability through offering dask as a service



AI in DeepESDL

- Dedicated Task for enhancing ESDL's **AI** capabilities
- Building on **Python's** extensive AI stack
- **GPU** profiles available
- Including tools for **tracking, managing, and archiving** experiments
- **Collaborative** research fostered by implementing **teams** with joint access to resources and services



DeepESDL persisted Datasets

Earth System
Data Cube
(ESDC) v3.0.1

Black Sea
Data Cube

Currently numerous
different analysis -
ready datacubes are
available via DeepESDL.
With more than 70 variables!

BICEP Pools and
Fluxes of the Ocean
Biological Carbon Pump

SMOS Freeze and
Thaw Processing
and Dissemination
Service

Land
Cover Maps from
ESA and C3S

Hydrology
cube

Polar Cube



<https://deepesdl.readthedocs.io/en/latest/datasets/datasets/>

DeepESDL Datastores

ESA CCI
Climate Change
Initiative

CMEMS
Copernicus
Marine Service

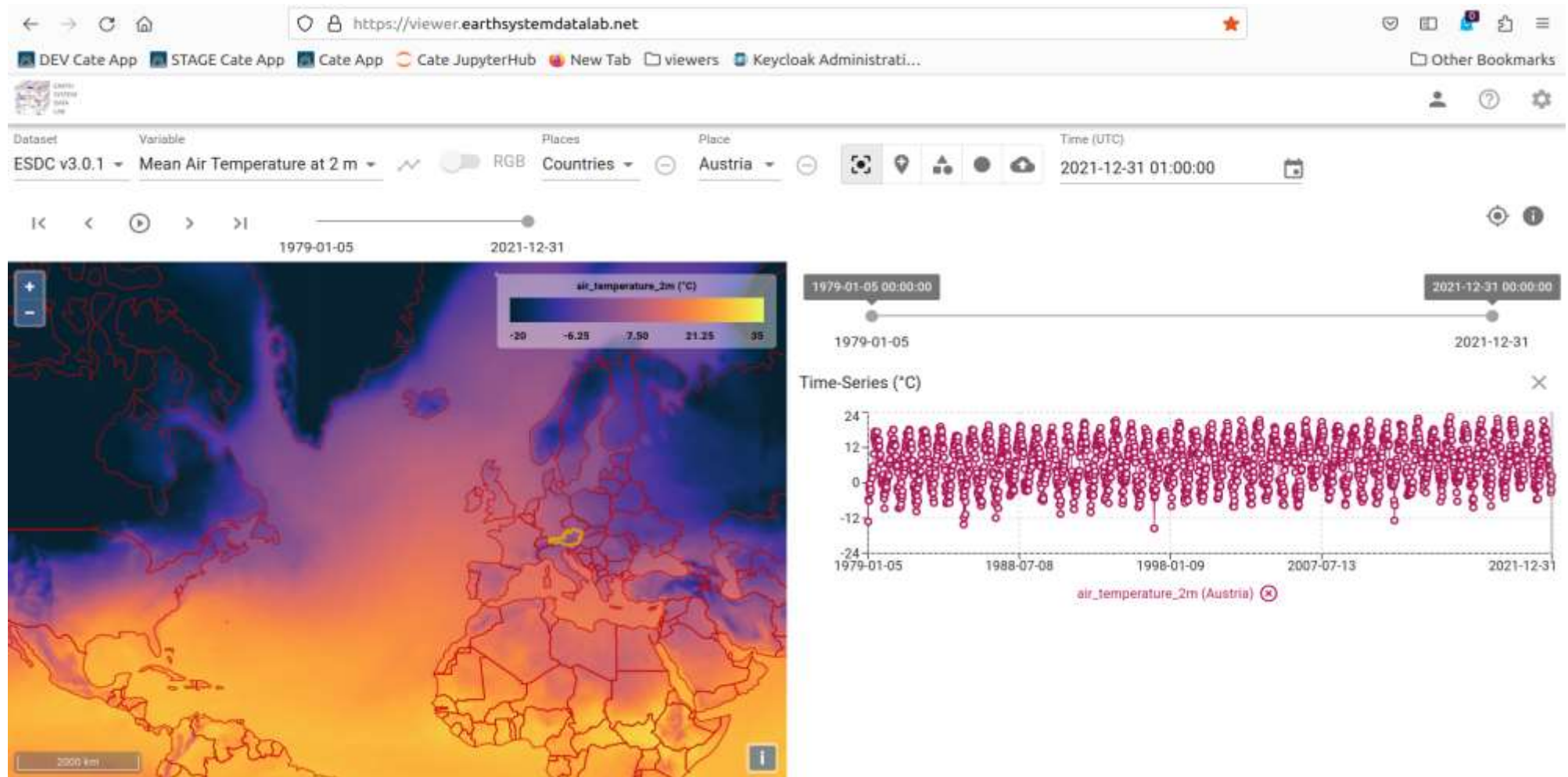
C3S
Copernicus Climate
Change Service

Sentinel Hub
Data Access API of
Copernicus DataSpace
Ecosystem

Here you have the overview of on the fly access to relevant data services from ESA and Copernicus.



DeepESDL xcube Viewer



Cube generation recipes

The screenshot shows the GitHub interface for the repository 'deepsdl/cube-gen'. The top navigation bar includes 'Pull requests', 'Issues', 'Codespaces', 'Marketplace', and 'Explore'. The repository name 'deepsdl/cube-gen' is displayed, along with 'Public' status, 'Edit Pins', 'Watch' (4), 'Fork' (0), and 'Star' (0) buttons. Below the repository name, there are tabs for 'Code', 'Issues' (18), 'Pull requests' (3), 'Discussions', 'Actions', 'Projects', 'Wiki', 'Security', 'Insights', and 'Settings'. The main content area shows a pull request by 'davemtz' titled 'FIX: Updated inputs-collect code and recipe to account for additional...'. Below the pull request, there is a file list for the 'Ocean' directory, including folders like 'inputs-collect', 'inputs-preprocess', 'notebooks', 'output-merge', and 'output-postprocess', and files like 'README.md', 'cube4km.geojson', and 'cube9km.geojson'. The right side of the screenshot shows the 'README.md' file content, which is titled 'Ocean Datacube Generation'. The README text includes: 'Information dedicated to the generation of the Ocean Datacubes. By using the following recipe two ocean data cubes can be reproduced.', a list of two data cubes (POC at 4 km and Oceanic Export Production at 9 km), the total storage required (490 GB), and a section titled 'Cubes generation process' which is divided into four phases. The first phase is '1. Downloading required raw datasets', which describes how datasets were downloaded into a 'data' folder and provides instructions on how to download them using a specific python environment.

<https://github.com/deepsdl/cube-gen>

19th February 2024

DeepESDL Introduction

11

DeepESDL in the Network of Resources

The screenshot shows the NoR web portal interface. At the top left is the NoR logo. To the right are navigation links: Service Discovery, About, and Web Portal Support. The main content area features the Brockmann Consult logo and the service name 'Brockmann Consult - DeepESDL'. Below this is a horizontal menu with buttons for Details, Collections, SLA, Provider Helpdesk, Training, and Pricing Wizard. Underneath the menu are two expandable sections: '+ Collections Overview' and '+ Service Offering Overview'.

The screenshot shows a form titled 'Ask ESA for Sponsorship'. The form contains the following elements:

- A text prompt: 'Please provide the information below before starting the application for ESA sponsorship.'
- A dropdown menu labeled 'ESA Project' with the value 'NO' selected.
- A dropdown menu labeled 'I will co-fund the cost' with the value '-- select an option --' selected.
- A note: 'Please note that after the submission of the sponsoring request, the data provided will be handled according to the [ESA Privacy Notice](#).'
- A button at the bottom labeled 'Ask ESA for Sponsorship'.

- DeepESDL services can be purchased via the Network of Resources
- Modular service offer
- ESA sponsorship available:

<https://nor-discover.cloudeco.group/Service/Brockmann-DeepESDL/SponsoringWizardPricelist>

Introduction to the demo

- For the demo, we'll use the notebooks in the team shared folder.
- For accessing the demo notebooks, kindly use the command below to create a symbolic link to the shared directory
In -s /extra ~/team-shared
- Several useful DeepESDL / xcube features are demonstrated:
 - Creation of atmospheric cubes using CDS store
 - Exploration of the data and metadata
 - Writing cubes to team storage
 - Visualize data cubes using xcube viewer jupyterlab extension
 - Using zappend tool to create zarr cubes from time slices
 - Sharing notebooks and files within the team