cloudflight

Fostering collaborative working

by means of notebook development platforms and connected services in the field of large aerosol and cloud data processing

- an OLCI data application example

ATMOS 2024, 2024-07-04

Patrik Kovacs & Verena Lanzinger

Agenda

- Collaborative Cloud Working
- Tech Stack
- Example Use Cases
- Demo

01

Collaborative Cloud Working Status Quo

Current challenges

- General **notebook environment** on cloud with limited or no **collaboration options**
- **EODATA** availability
- Plugging **custom data** into the environment
- **Python environments** are built and shared manually by each user
- Dealing with cloud dashboards to set up a **parallel processing**

Collaborative working and joint development solutions

- ❑ Customization (S3 LSR) (DIVA)
- ❑ Kubernetes (S3 LSR) (DIVA)
- ❑ (DASK) Parallelization (S3 LSR)
- ❑ EO data on modern DIAS (S3 LSR)
- ❑ Shared environments (e.g., conda) (S3 LSR) (DIVA)
- ❑ "Bring your own data" solution (S3 LSR)
- ❑ No additional step of transferring the results (S3 LSR) (DIVA)
- ❑ Collaborative working – joint development (S3 LSR) (DIVA)
- ❑ User management – permission (S3 LSR) (DIVA)

02

Technology – Tech Stack

03

Use case: S3 LSR

Use Case Sentinel-3 Land Surface Reflectivity Auxiliary Product

EUMETSAT funded project

Project Partners

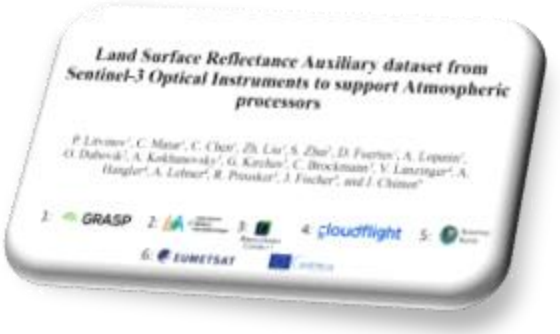
GRASP · Brockmann Consult · EUMETSAT · Spectral Earth · Copernicus · cloudflight

Motivation

Creating OLCI global product with a highly accurate retrieval algorithm using EODATA and auxiliary data as input.

Challenge	Solution
Implementing a prototype on a remote environment	Collaborative notebook environment in the cloud with share space
Ensuring data availability	Accessing EODATA and plugging in the auxiliary/custom input data (e.g. cloud-mask)
Team members are using the same Python environment	Building Python environments in a shared environment manager
Establishing the global product with the implemented prototype	Creating a parallel cluster managed directly from the notebook environment and process the global input on the cluster running the code in the cloud
Sharing the results	Directly sharing the output storage with the community/partners

Presented by
Dr. Pavel Litvinov at the Sentinel 3 Validation Team Meeting in Darmstadt December 2023



SCAN to ABSTRACT

04

Use case: DIVA

Use Case DIVA - Demonstration of an **I**ntegrated approach for the **V**alidation and exploitation of **A**tmospheric missions

ESA funded project

Project Partners

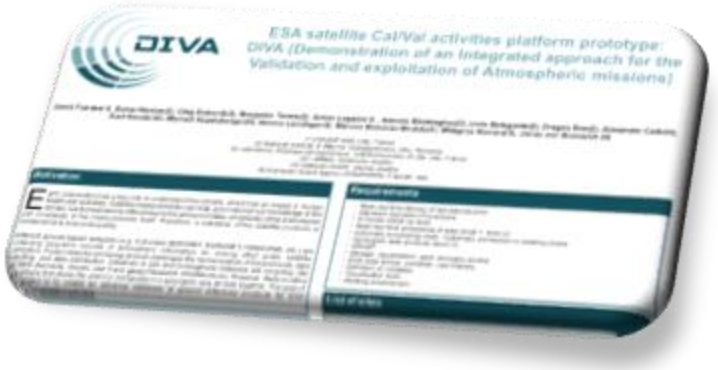
LUFTBLICK EARTH OBSERVATION TECHNOLOGIES · GRASP · IDOC 2000 · esa · cloudflight

Motivation

Creating a custom platform for specific data collected from different sources (Ceilometer, Spectrometer, Lidar, and Photometer), tailored to the needs of the user groups

Challenge	Solution
Collaborative notebook environment for different scientific groups with shared space	Jupyter environment with shared data space and shared Python environments
Customized user API for accessing the project specific data	Customized versioned Python packages installed automatically into the environment
User and permission management of these group members	Permission handling with Keycloak

Presented by Dr. David Fuertes at the ATMOS conference in 2021

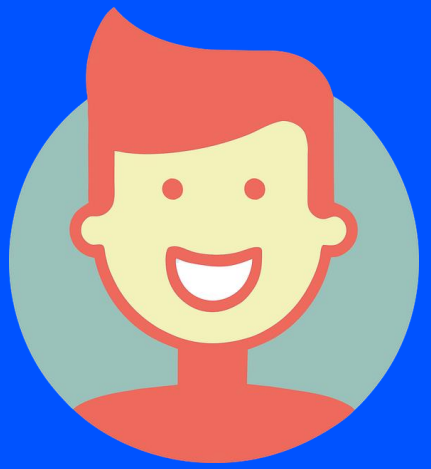


SCAN to POSTER

05

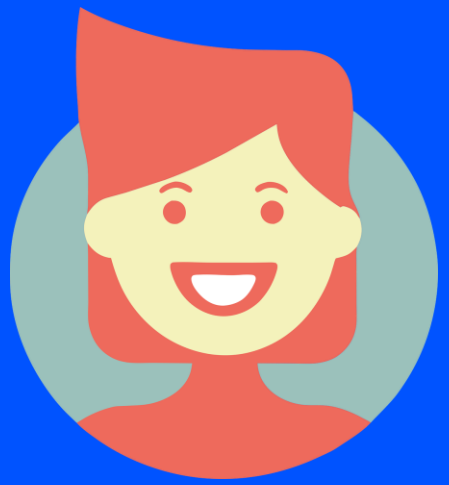
Demo

Let‘s talk about solutions !



Patrik Kovacs

patrik.kovacs@cloudflight.io



Verena Lanzinger

verena.lanzinger@cloudflight.io