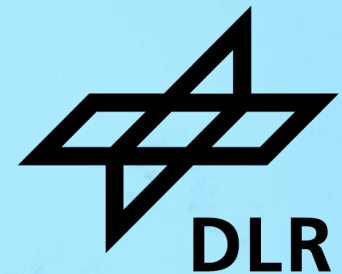


# terrabyte offerings

Dr. Jonas Eberle, Max Schwinger, terrabyte team  
January 2024



# Agenda



- Identity Management
- HPC
- Storage & Data Services
- Compute Portal
- Processing Services
- Support

# terrabyte a high-performance data analytics (HPDA) platform



**June '23:**  
Official launch of the  
terrabyte infrastructure  
at LRZ.

**July '23:**  
Launch of terrabyte  
web services



# terrabyte – status quo

## Hardware



### Data storage (LRZ)

- 50 PB net
- Currently 6 PB dedicated for users
- GPFS Filesystem

### Compute environment (LRZ)

- ca. 44.000 CPU vCores with 273 TB RAM and 188 GPUs with 47 TB RAM + 14 TB GPU RAM



### Network line between OP and LRZ

- 10 Gbit/s line (upgrade to 100 Gbit/s possible)

### Internet line between LRZ and external hubs

- 20 – 40 Gbit/s to Copernicus Data Space Ecosystem
- 10 Gbit/s to NASA Alaska Satellite Facility

- Sentinel-1: SLC, GRD, NRB
- Sentinel-2: L1C, L2A
- Sentinel-3
- Sentinel-5p
- Landsat: Level-2
- MODIS: MOD09GQ, MOD10A1, MOD13A2, MOD13A3, MOD13Q1  
MYD09GQ, MYD10A1, MYD13A2, MYD13A3, MYD13Q1
- VIIRS/NPP: VNP09GA, VNP13A1, VNP15A2H
- Meteosat: SEVIRI L1.5 DISC / HRV / VISIR
- DEM: Copernicus DEM 30m, 90m  
SRTM DEM, Aster DEM

# terrabyte Offerings

An overview



Jira Service Desk

Discourse

terrabyte Support

jupyter QGIS R VS Code

terrabyte Compute Portal

open EO Open Geospatial Consortium

terrabyte Processing services\*

lrz

Authentication Register IDM-Portal 2FA-Portal

terrabyte Identity Mng.

slurm workload manager

terrabyte HPC

DSS Web Globus Online

terrabyte Storage services

eo API STAC

terrabyte Data services

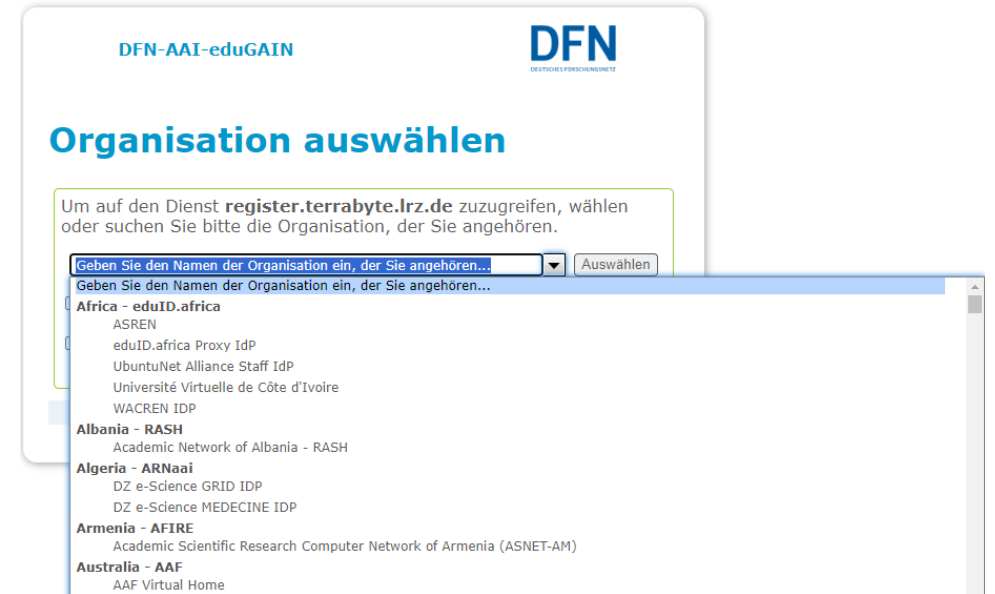
\* In development

# terabyte Offerings

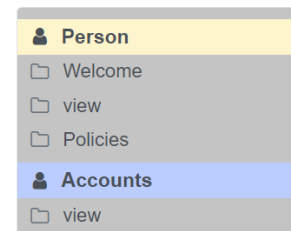
## Identity Management



- <https://register.terabyte.lrz.de>
  - Create a LRZ account based on your DLR login (or other research organizations)
  - All accounts from DLR login have access to terabyte automatically
  - External accounts need to be added to a project



- <https://idmportal2.lrz.de>
  - User Management (e.g., password reset, user groups)
  - User Policies (e.g., export control regulations)
  - Management of "projects"



[> Logout



### Identity Management Portal

Account: di93jov - Benutzer: Eberle, Jonas

In case of problems with or questions about your accounts, please contact the project manager or the responsible of the corresponding account.

#### Personal accounts

Account	Status	Creation	Expiration	UID	GID
di93jov	active	Jan 19, 2023		██████	██████

# terabyte Offerings

## Identity Management



- <https://simmfa.sim.lrz.de>
  - Create and manage second factor authentication (2FA) for terrabyte services
  - Supports multiple 2FA devices: mobile device, Yubikey, TAN list
  - <https://docs.terabyte.lrz.de/services/identity/setup-2fa/>
- When do I need a LRZ project?
  - Work with non-DLR users (e.g., external project partners)
  - Internal project with multiple DLR institutes involved and dedicated storage capacity is necessary.
  - <https://docs.terabyte.lrz.de/services/project-management/introduction/>

Recommended Tokens for Login Procedures	TOTP Yubikey
Recommended Recovery Token	TAN List
Additional Token	PUSH

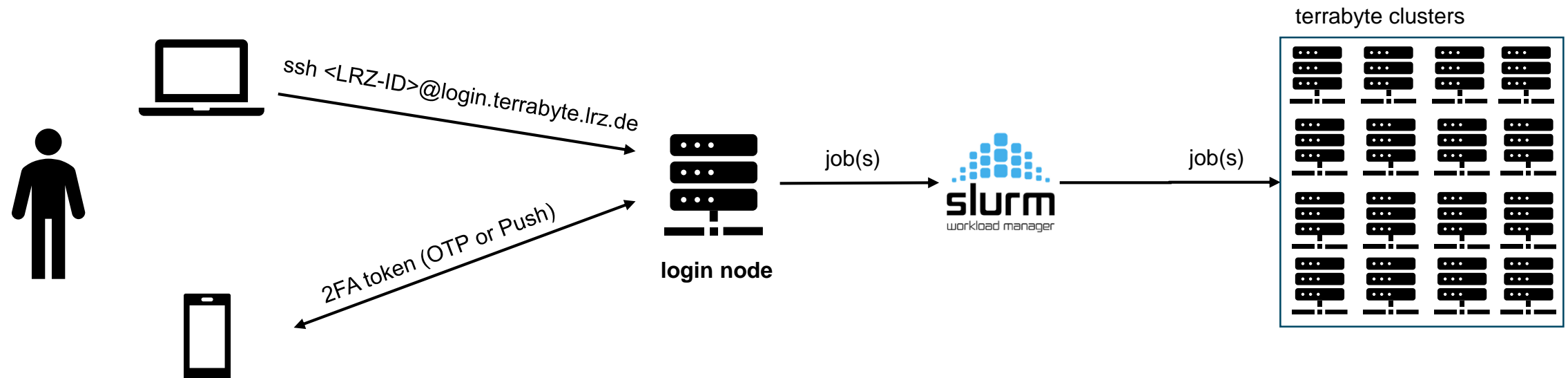


# terrabyte Offerings

## HPC



- Console login
  - With password authentication: **login.terrabyte.lrz.de**
  - With ssh key authentication: **login-keyauth.terrabyte.lrz.de**



# terabyte Offerings

## Storage Services



- Data Science Storage: <https://docs.terabyte.lrz.de/services/data-storage/introduction/>

Purpose/what to store here	Availability	Where to access the files	Space Available	Automatic backup
<b>Home-Directory</b>				
User's Home Directory Important data like scripts or personal data	By default Read-write access	\$HOME or /dss/dsshhome1/lxc##/<user id>	100 GB <b>per user</b>	<b>YES</b> , file system snapshots  For details see Snapshots
<b>Data Science Storage (DSS): terrabyte Geodata</b>				
DSS (Data Science Storage) All available Earth observation data sets and auxiliary data on HPDA terrabyte are provided here	By default <b>Read access only!</b>	<b>!STAC-Metadata Catalog!</b> or /dss/dsstbyfs01/pn56su/pn56su-dss-0008 /<DEM, Landsat, Meteosat, MODIS, Sentinel-1, Sentinel-2, Sentinel-5p>	-	-
<b>Data Science Storage (DSS): Project-specific storage</b>				
DSS (Data Science Storage) Private (single user) or project (multiple users) storage for e.g. results, individual input data, ...	<b>Upon request</b> , for details see Request project-specific storage Read-write access	/dss/dsstbyfs02/<project-id>/<storage-container-id>/<user-id>  View and select an appropriate directory and define your own variable (e.g. WORK) in your (e.g. ~/.profile or ~/.bashrc)	Individual  In general, a few TB per user/project are no problem.	<b>NO</b>

# terabyte Offerings

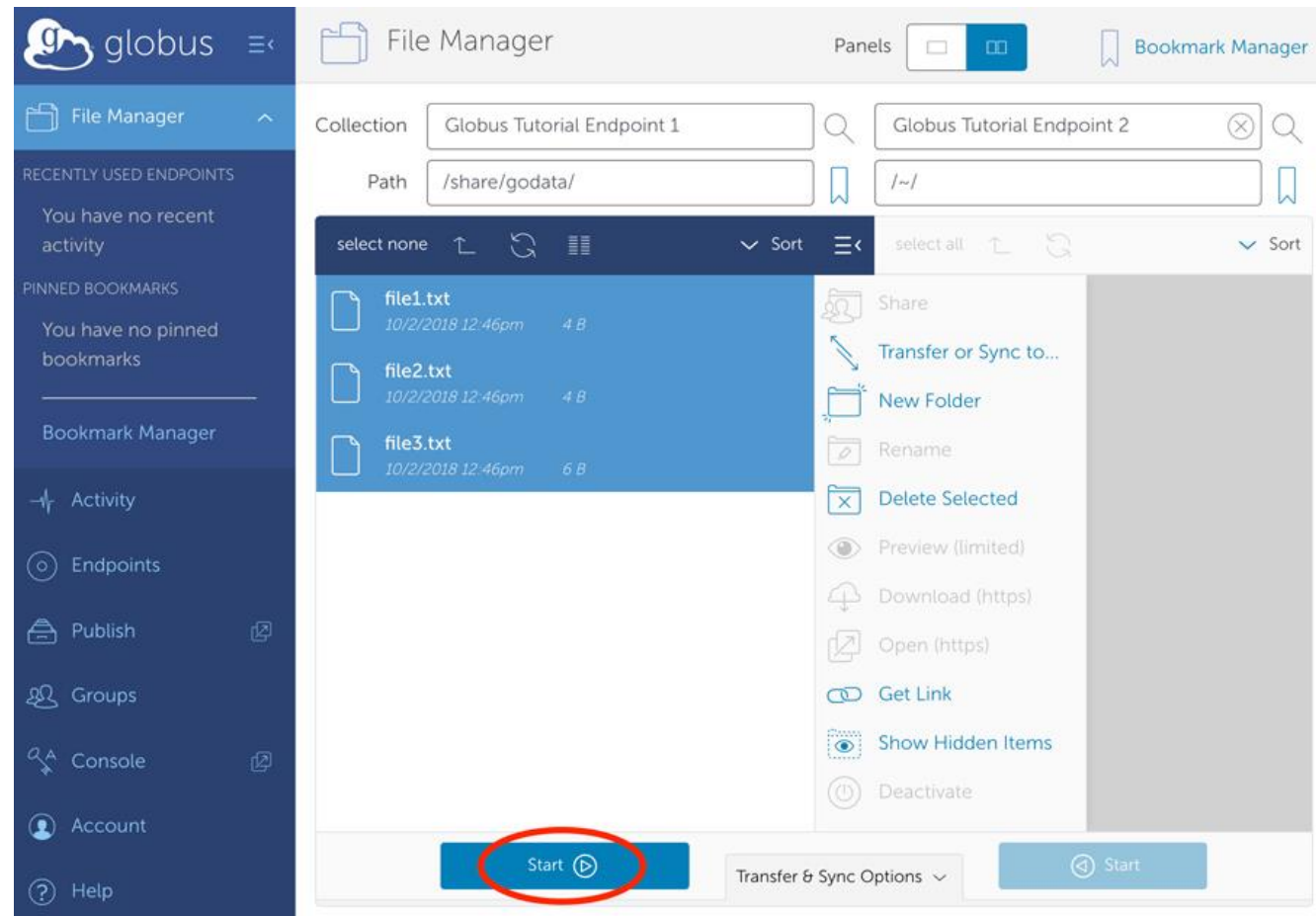
## Storage Services



### ■ Globus Online

- Data transfer between two Globus Online endpoints
- All DSS containers can be accessed through Globus Online
- Users need to setup an own Globus Online endpoint to transfer data
- Data can be shared with non-terabyte users
- Web (<https://globus.org>) and command line interface

<https://docs.terabyte.lrz.de/category/large-scale/>



# terabyte Offerings

## Data Services



- SpatioTemporal Asset Catalog (STAC)
  - Metadata specification for catalogs, collections, and items
- STAC Application Programming Interface (API)
  - <https://stac.terabyte.lrz.de/public/api>
  - Filter and search metadata catalog
  - Soon: Private user and project catalog
- Browser
  - <https://stac.terabyte.lrz.de/browser>



### terabyte STAC API

API Source Share Language: English

Browse Search

#### Description

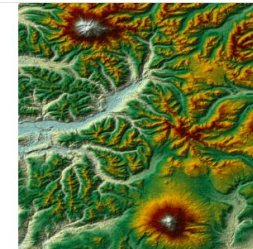
Curated data catalog of the terabyte platform available at <https://stac.terabyte.lrz.de/public/api>

#### Additional Resources

- OpenAPI service description
- OpenAPI service documentation

Catalogs 20 Tiles List Ascending Descending

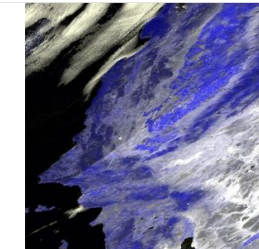
Filter catalogs by title



Copernicus DEM GLO-30

**GeoTIFF** The Copernicus DEM is a digital surface model (DSM), which represents the surface of the Earth including buildings, infrastructure, an...

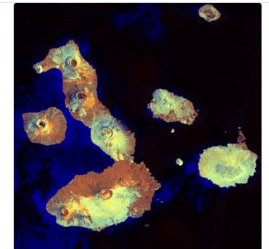
4/22/2021, 12:00:00 AM UTC



MOD09GQ.061 & MYD09GQ.061:  
MODIS Surface Reflectance Daily  
(250m)

**Hdf4 HDF** The MOD09GQ/MYD09GQ Version 6.1 product provides an estimate of the surface spectral reflectance of Terra/Aqua Moderate...

2/24/2000, 12:00:00 AM UTC - 6/10/2023, 12:00:00 AM UTC



Sentinel-1 SLC Level-1

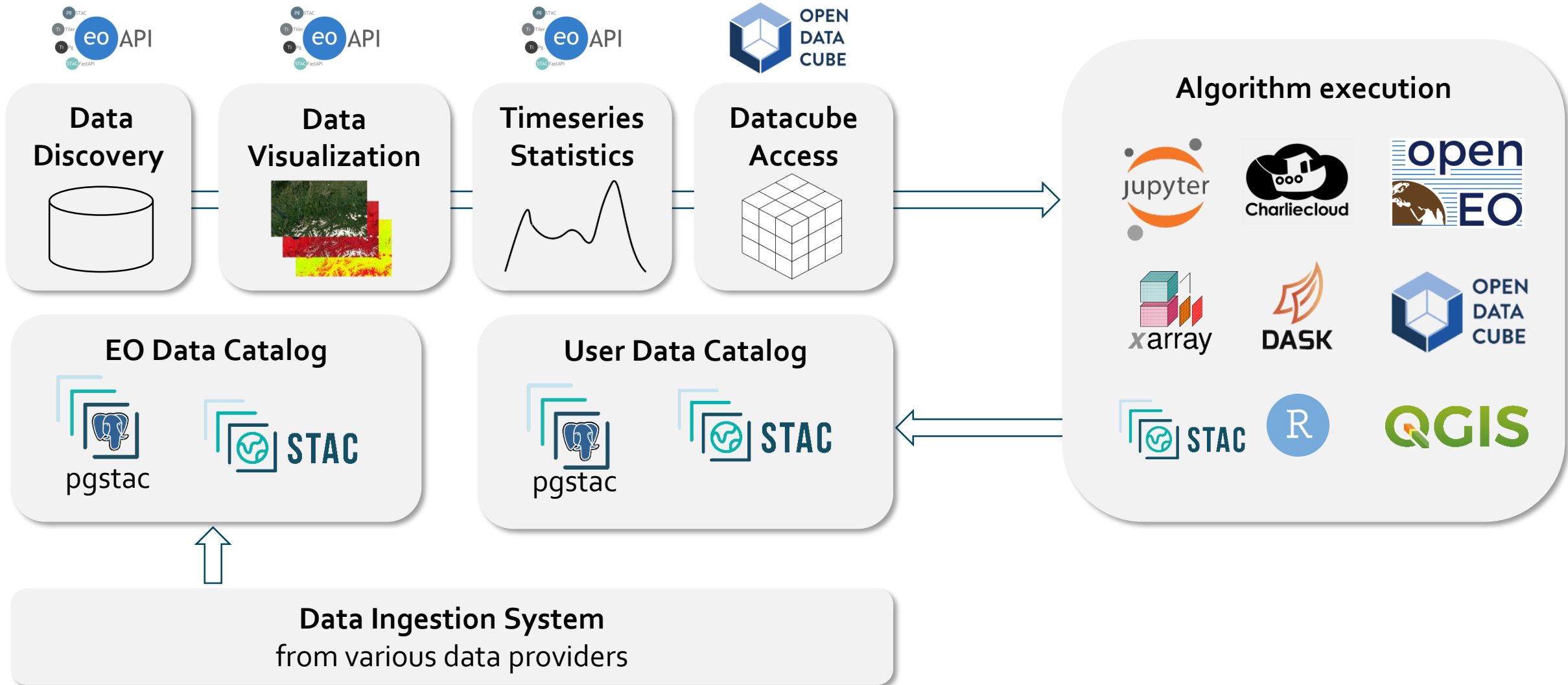
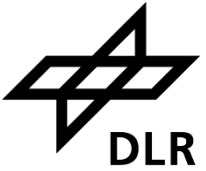
This is the original Sentinel-1 SLC Level-1 dataset provided from ESA.

10/3/2014, 3:08:21 PM UTC - 6/14/2023, 3:45:08 PM UTC



# terabyte Offerings

## Data Services

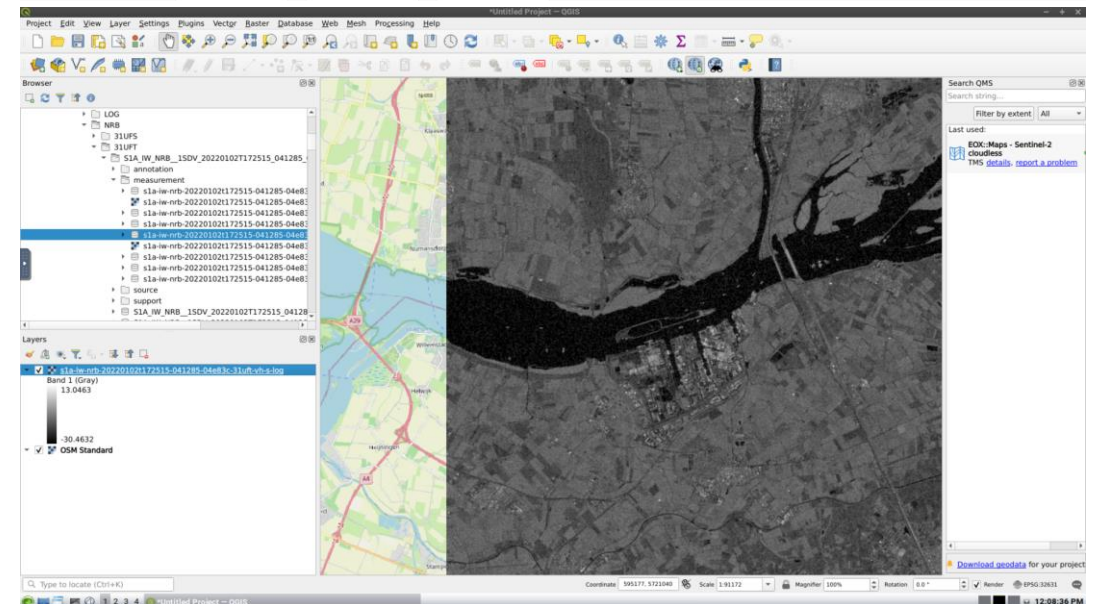
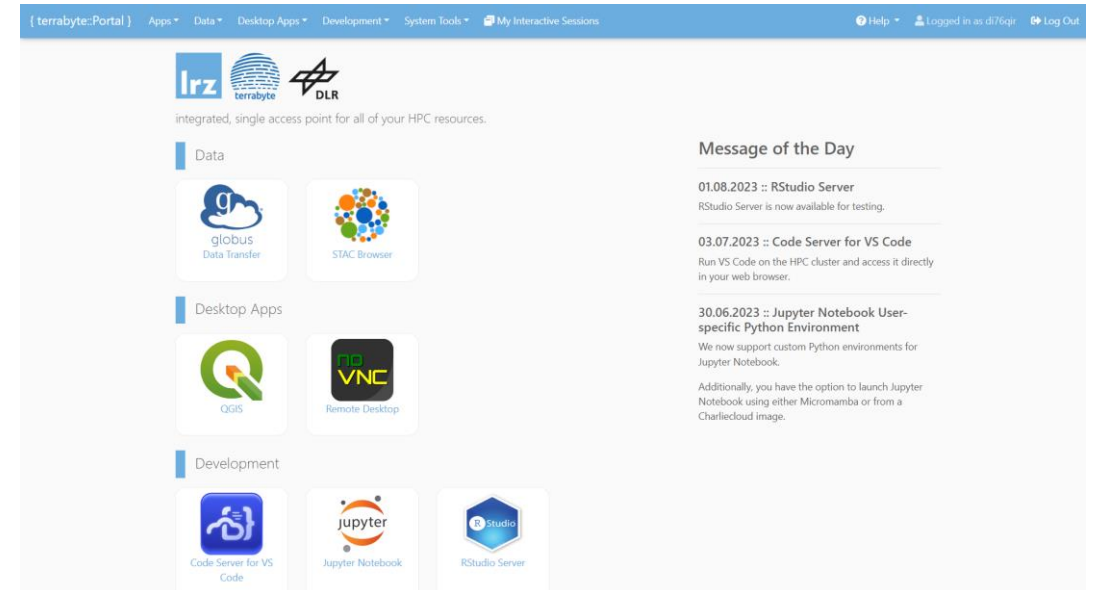


# terrabyte Offerings

## Compute Portal



- <https://portal.terrabyte.lrz.de>
  - Operated by LRZ
  - List, monitor, and manage your HPC jobs
  - Start/Stop interactive applications
    - Jupyter Notebook/Lab
    - QGIS in the browser
    - RStudio Server
    - Visual Studio Code



# terabyte Offerings

## Processing Services



- Objective: Make algorithms/processors available for on-demand execution by others (without sharing source code)
- Considerations
  - ESA ITTs often require algorithms to be developed with reproducible technologies
    - OpenEO
    - OGC EO Application Package
  - Run algorithms on other platforms (e.g., CODE-DE)
- Technologies under investigation for terrabyte
  - OpenEO API
  - OGC Application Deployment & Execution Service for OGC EO Application Package

# terrabyte Offerings

## Support



- Web: <https://support.terrabyte.lrz.de>
  - Forum for various questions & comments
  - Engagement of all users is welcome!
  - Share knowledge with all users.
- Email: [servicedesk@terrabyte.lrz.de](mailto:servicedesk@terrabyte.lrz.de)
  - JIRA-based ticketing system (e.g., for more private issues)
- Mailing list: [terrabyte@lists.lrz.de](mailto:terrabyte@lists.lrz.de)
  - Used for announcements and information about maintenance via mail
  - Join by sending email to [terrabyte-join@lists.lrz.de](mailto:terrabyte-join@lists.lrz.de)

A screenshot of the terrabyte support forum interface. On the left is a navigation sidebar with categories like Topics, My Posts, Review, Admin, More, CATEGORIES (Announcements, General, terrabyte Portal, HPC Slurm, STAC API, Curated data, Data requests, Feedback), TAGS (sentinel-2, All tags), MESSAGES (Inbox), CHANNELS (General, Staff), and PERSONAL CHAT. The main content area shows a list of forum categories with their respective topic counts and descriptions. The categories are: Announcements (0 topics), General (1 topic: 'Welcome to our community! 17d'), terrabyte Portal (0 topics), HPC Slurm (0 topics), STAC API (0 topics), Curated data (0 topics), Data requests (2 topics: 'MODIS/VIIRS Cloud Optimized Geotiff 6d' and 'Sentinel-2 Collection 1 L2A 6d'), and Feedback (0 topics).

Category	Topics	Latest
<b>Announcements</b> We will post updates and downtimes (e.g., for maintenance) in this category.	0	
<b>General</b> Create topics here that don't fit into any other existing category.	1	* Welcome to our community! 17d
<b>terrabyte Portal</b> Discussions and questions about the terrabyte portal available under <a href="https://portal.terrabyte.lrz.de">https://portal.terrabyte.lrz.de</a>	0	
<b>HPC Slurm</b> Discussions and questions about the usage of the terrabyte HPC Slurm environment.	0	
<b>STAC API</b> Please ask questions about the use of the STAC API available at <a href="https://stac.terrabyte.lrz.de/public/api">https://stac.terrabyte.lrz.de/public/api</a> .	0	
<b>Curated data</b> Discussions and questions about the curated data provided on terrabyte. For new data requests, please see the other category!	0	
<b>Data requests</b> If you need data, which should be provided as part of the curated data catalog, please open a topic under this category. Please always search for existing topics!	2	MODIS/VIIRS Cloud Optimized Geotiff 6d Sentinel-2 Collection 1 L2A 6d
<b>Feedback</b> General discussion about the terrabyte platform and its services, how it works, and how we can improve it.	0	



# terrabyte Offerings Support

<https://github.com/DLR-terrabyte/eo-examples>



## Examples:

The screenshot shows the GitHub repository page for `DLR-terrabyte / eo-examples`. The repository is public and has 2 stars and 0 forks. The main branch is selected, and there is 1 branch and 0 tags. The commit history shows a recent commit by Jonas Mueller on August 1st, adding reprojction, scaling and masking for MODIS notebook. The file list includes `datasets`, `quickstarts`, `.gitignore`, `LICENSE`, and `README.md`. The README content is displayed below, starting with the title "Examples of working with Earth observation data on terrabyte".

File	Description	Time
<code>datasets</code>	Added reprojction, scaling and masking for MODIS notebook.	4 months ago
<code>quickstarts</code>	Initial examples collected from the terrabyte team	4 months ago
<code>.gitignore</code>	Added reprojction, scaling and masking for MODIS notebook.	4 months ago
<code>LICENSE</code>	Initial commit	4 months ago
<code>README.md</code>	Update README.md	4 months ago

### Examples of working with Earth observation data on terrabyte

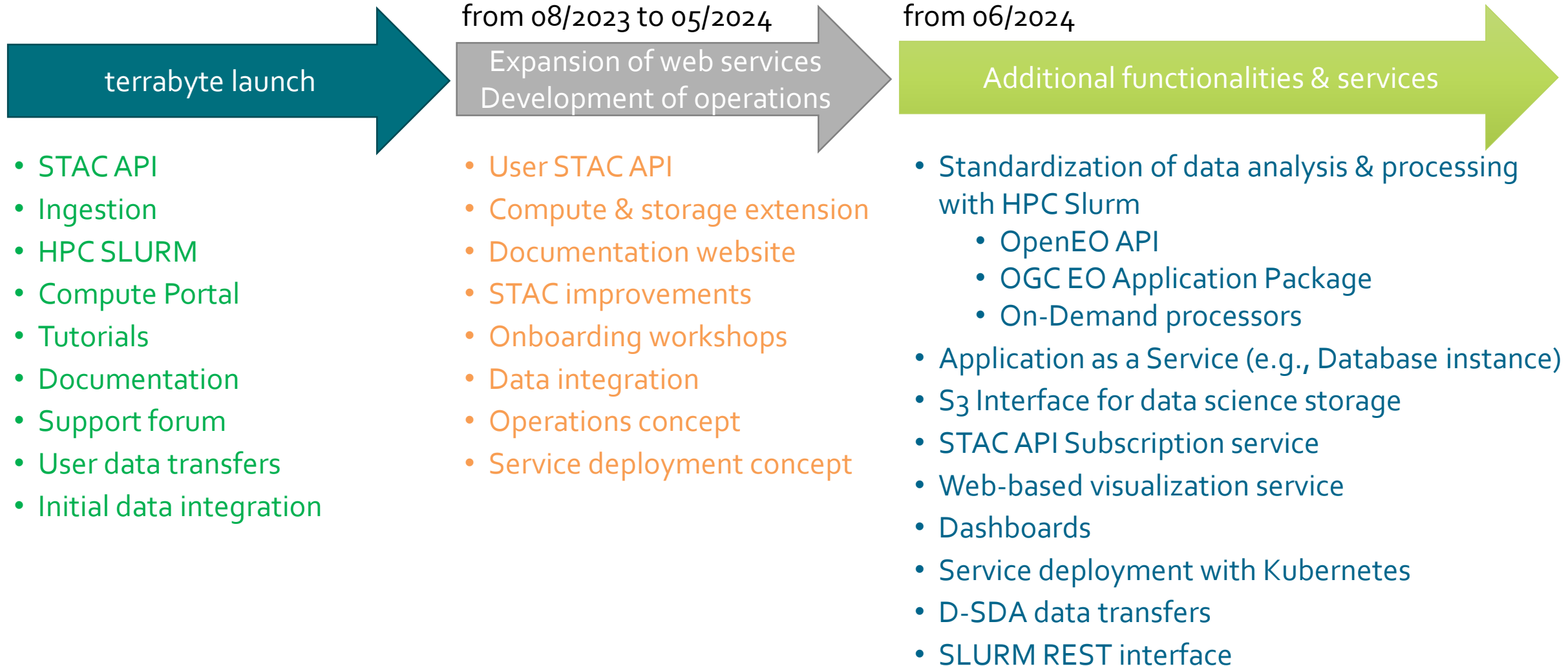
The terrabyte Team provides initially basic examples on how to work with Earth Observation data on the terrabyte Platform.

Everyone is welcome to contribute with own examples (please open a pull request).

Issues with the notebooks can be discussed in the internal terrabyte user- and support forum.

### SpatioTemporal Asset Catalog (STAC)

# Roadmap



Thank you for your attention!

Project coordination



**Maximilian Schwinger**

DLR DFD-IBS

Phone: +49 8153 28 1418

Email: [Maximilian.Schwinger@dlr.de](mailto:Maximilian.Schwinger@dlr.de)



**Dr. Jonas Eberle**

DLR DFD-INF

Phone: +49 8153 28 3495

Email: [Jonas.Eberle@dlr.de](mailto:Jonas.Eberle@dlr.de)