

UNIVERSITY OF TWENTE.



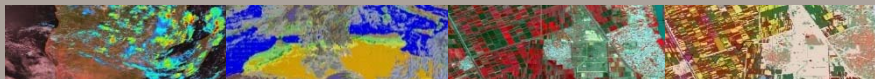
GEOCLIC SUMMER SCHOOL:

“ALL REQUIRED IS AN INTERNET CONNECTION”.

GEO-INFORMATION PROCESSING IN THE CLOUD
USING THE NEWLY ESTABLISHED COPERNICUS
DATA SPACE ECOSYSTEM.

BEN MAATHUIS
FACULTY ITC, UNIVERSITY TWENTE
THE NETHERLANDS

b.h.p.maathuis@utwente.nl



FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION





USE OF GEO-INFORMATION IN A CHANGING WORLD

Recap from Constanta workshop:

Nowadays:

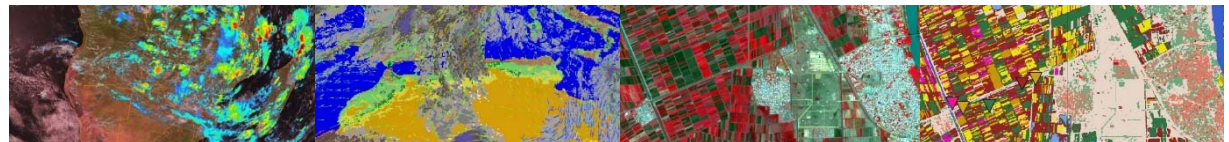
- Continued exponential growth of big EO data
- Technological advancement

Result in fundamental changes how EO data is stored and how users access, process and use the data:

From 'data to software' – to – 'software to data'



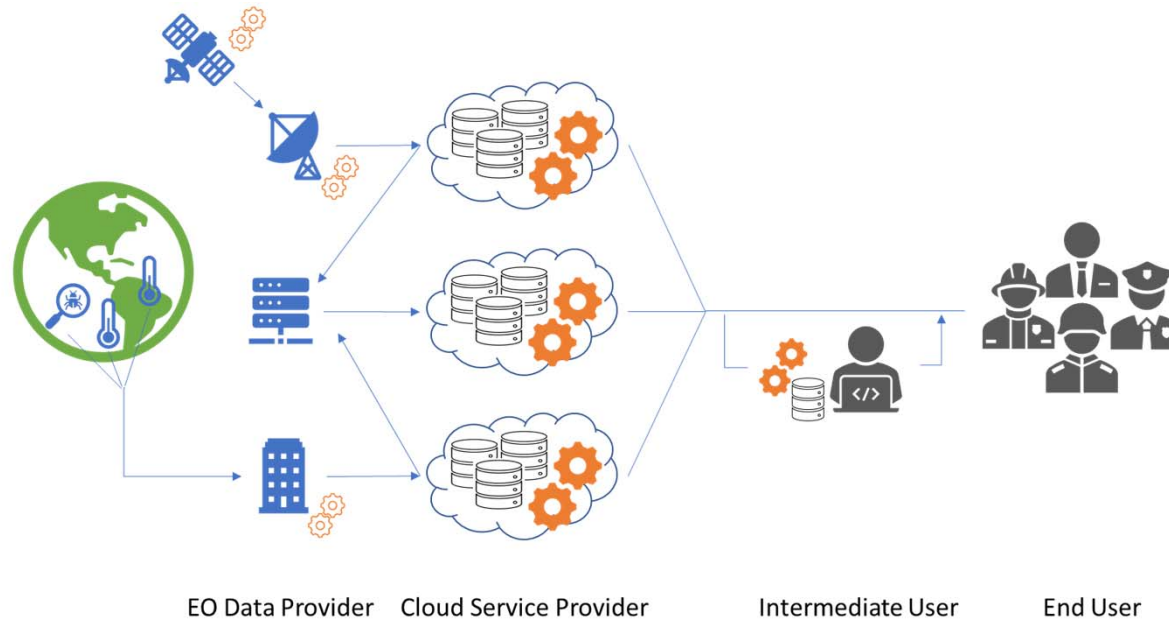
UNIVERSITY OF TWENTE.



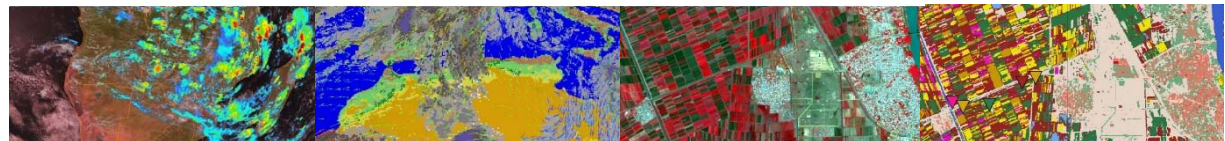


CLOUD-BASED DATA ACCESS AND PROCESSING

Instead of downloading EO data and processing it on local machines, intermediate users specify a virtual machine in the cloud and with this virtual machine, they access and process the data and only transfer intermediate or end results to their local machine.



UNIVERSITY OF TWENTE.



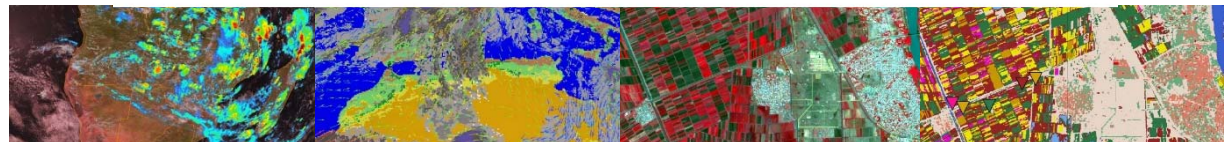
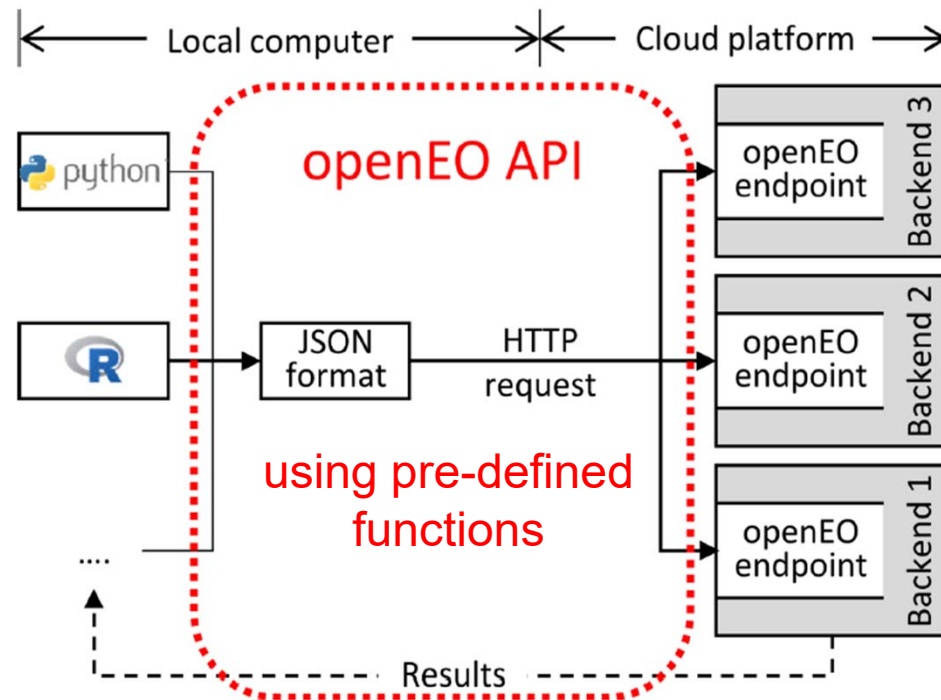


OPEN EO

- The open source 'OpenEO' API was developed offering possibilities of a standardised communication and of hiding architectural differences between cloud platforms, it reduces the described lack of cross cloud platform comparability.



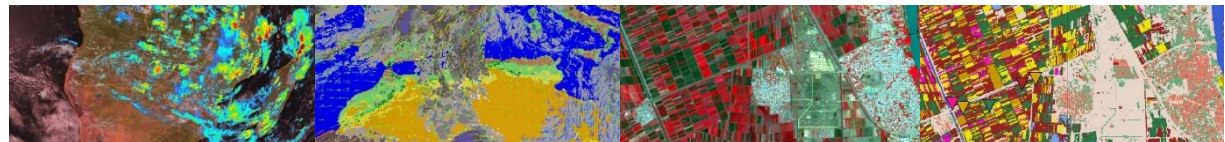
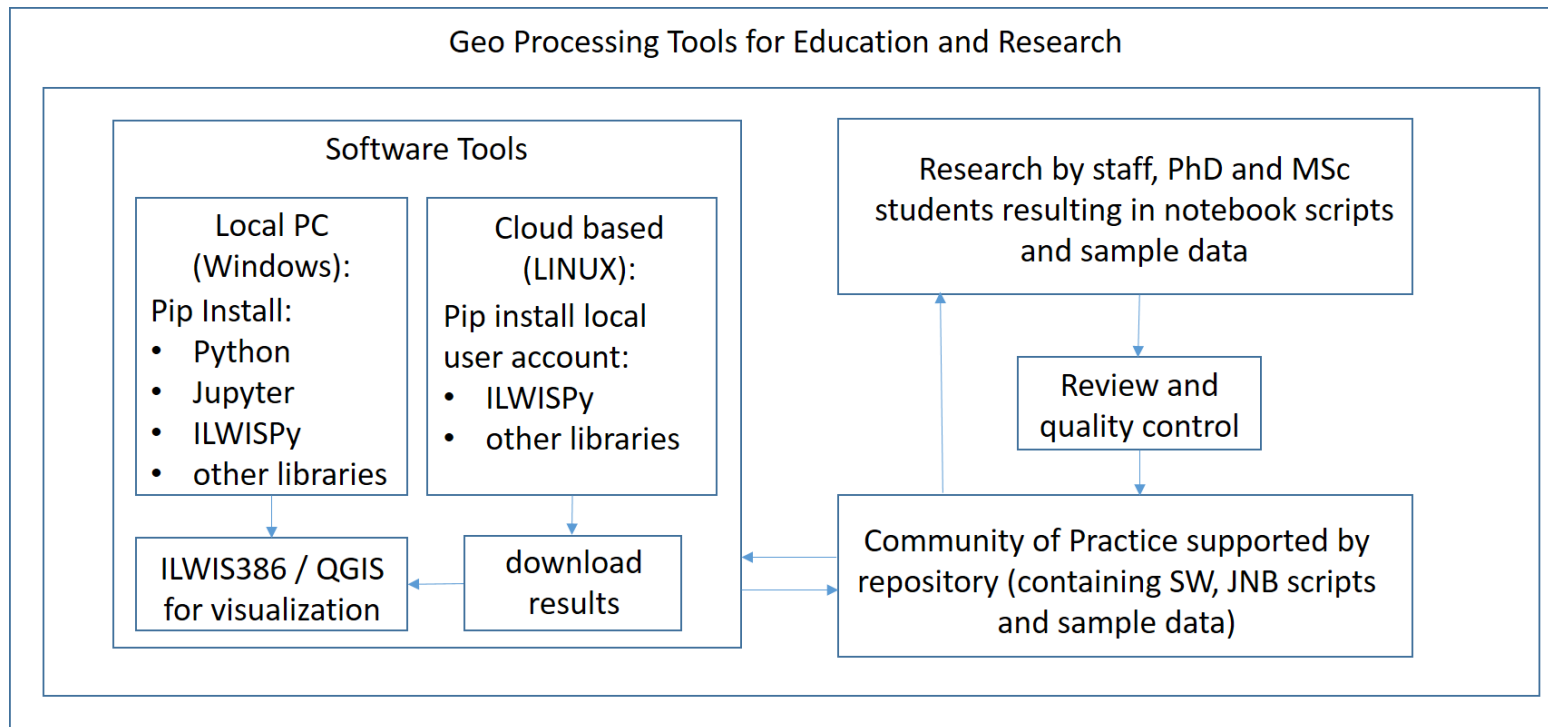
open API to connect Python to big Earth observation cloud back-ends in a simple and unified way, like Copernicus, GEE, VITO, etc.





EFFORTS UNDER WAY AT ITC

Develop a Community of Practice in Geo-information processing

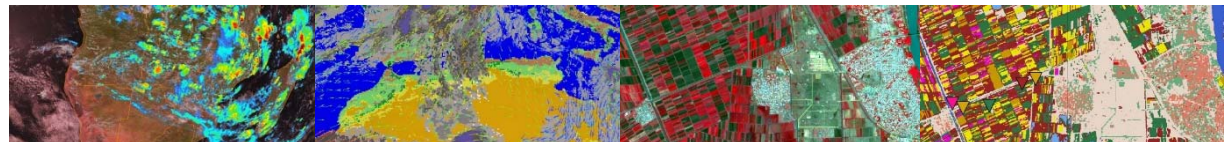




REPERCUSSIONS FOR STAFF / STUDENTS

- Getting (more) familiar with scripting languages:
 - from a graphical user interface to coding
 - time required to learn a scripting language
- Getting familiar with the various cloud based data resources and their access (using an application programming interface - API)
- New teaching approach: use of Jupyter Notebooks (code and markdown)
- Use of Python (multiple) site-packages
- From Windows OS to Linux

- Establishing 'your' cloud processing server??





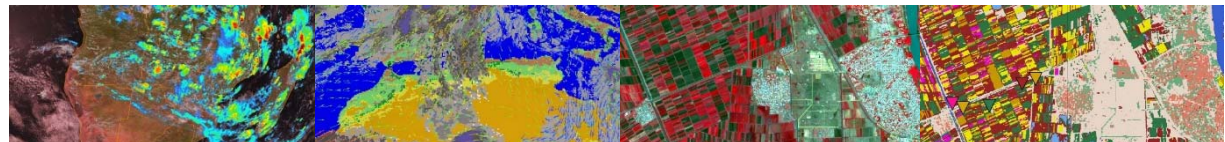
NEW DEVELOPMENTS

- Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu/>)

The screenshot shows the Copernicus Data Space Ecosystem website. The header includes the European Union logo, Copernicus logo, and CESA logo. Navigation links include News, Events, Gallery, Videos, and About. A main menu has 'EXPLORE DATA', 'ANALYSE DATA', and 'ECOSYSTEM'. A dropdown menu under 'ANALYSE DATA' lists 'APIs', 'Data Workspace', 'Traceability', 'Jupyter Notebooks', and 'openEO'. A 'SUPPORT' button and a 'MY ACCOUNT' button are also visible. The main content area features the heading 'Explore the Copernicus Data Space Ecosystem' and a welcome message: 'Welcome to the Copernicus Data Space Ecosystem, an open ecosystem that provides free instant access to a wide range of data and services from the Copernicus Sentinel missions and more on our planet's land, oceans and atmosphere. The Copernicus Data Space Ecosystem not only ensures the continuity of the open and free access to Copernicus data but also extends the portfolio for data processing and data access possibilities.' A 'LOGIN' button is at the bottom left. The URL bar shows 'https://dataspace.copernicus.eu/analyse/jupyter-notebooks'. The background features a satellite image of a coastal area with a data overlay.



UNIVERSITY OF TWENTE.

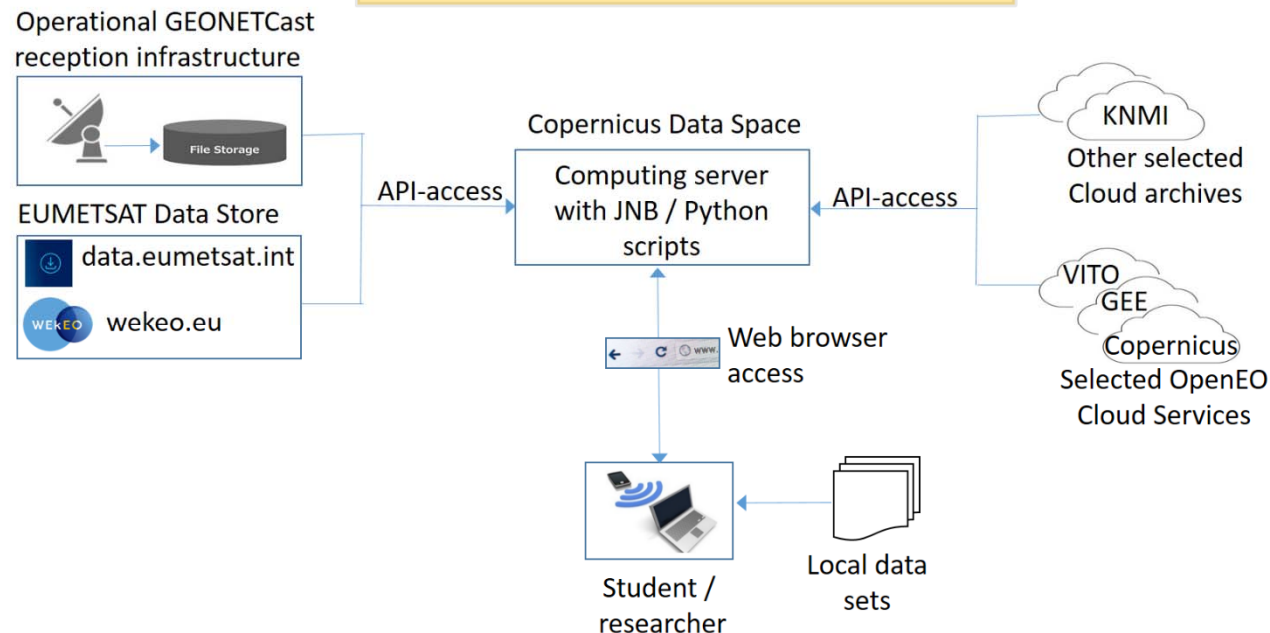




TEACHING GEO-INFORMATION PROCESSING IN THE CLOUD USING JUPYTER NOTEBOOKS

- The only requirement now is an internet connection!

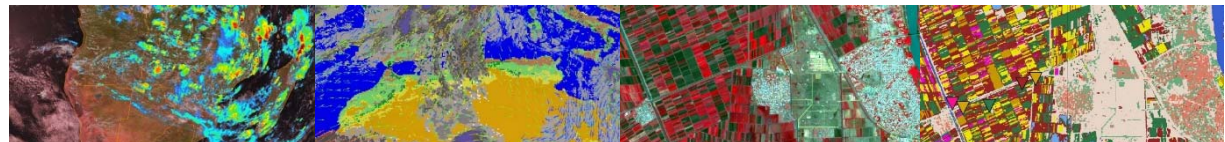
<https://dataspace.copernicus.eu/>



<https://documentation.dataspace.copernicus.eu/Applications/JupyterHub.html>



UNIVERSITY OF TWENTE.





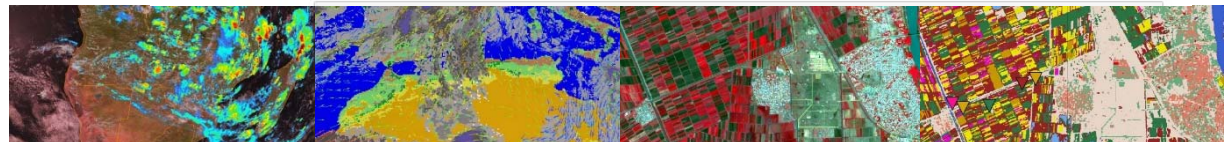
WHY USING JUPYTER NOTEBOOKS

- They're great for showcasing your training materials. You can see both the code and the results, export to pdf or HTML;
- Easy to use by students as a starting point. You can run cell by cell to better understand what the code does, supplemented by markdown cells, providing further explanation;
- Easy to host on server side and small file volume.

```
#### Create a loop - as markdown text

In [2]: #code cell with loop
for i in range(4):
    print(i)

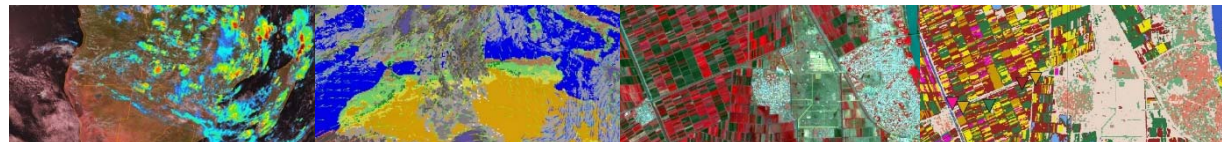
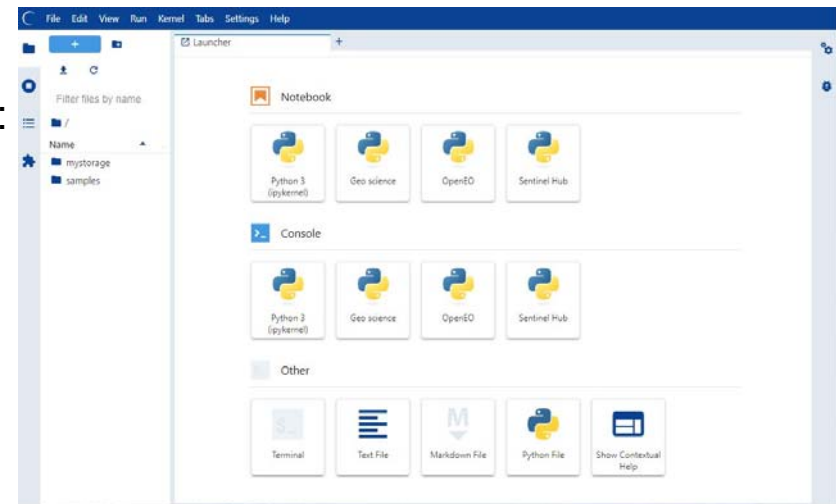
0
1
2
3
```





RESOURCES AT YOUR DISPOSAL

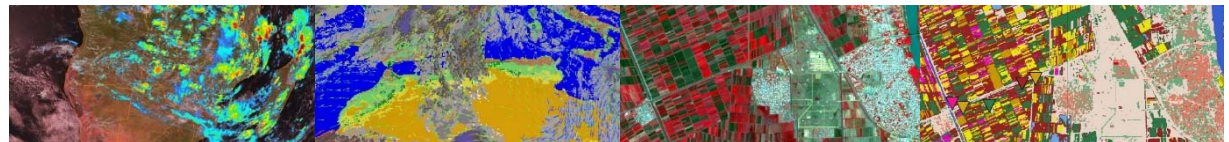
- Copernicus Data Space ecosystem available recently (July 2023)
 - Free access – easy registration
 - Direct access to Sentinel suite of data – full archive
 - In near future also data from contributing missions
 - Free usage of infrastructure provided:
 - Jupyter notebook support
 - Upon first login 2 folder are created:
 - ‘Mystorage’ (persistent–10 GB)
 - ‘Samples’ (re-created)
 - The availability of all other folders will be limited to your current session only.





TRAINING PACKAGE PREPARED

- Based on experiences from workshop in Constanta a package was prepared using notebooks in the free cloud resources provided:
 - Introduction to Python and Matplotlib
 - Introduction to descriptive statistics
 - Use of Python in Hydrology
 - Use of ILWISPy for Geo-Information processing
 - Introduction ILWISPy
 - Introduction to Remote Sensing using ILWISPy
 - Advanced Image Classification
 - Introduction OpenEO using ILWISPy
 - Introduction TimeSeries processing using ILWISpy
 - Introduction Satpy and use of data retrieved in ILWISPy

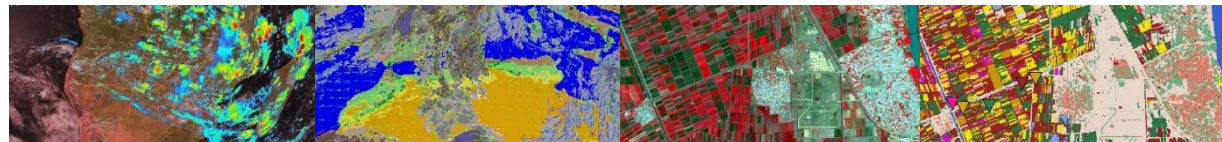
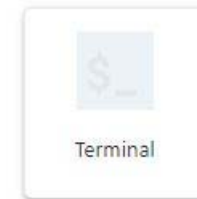




TRAINING PACKAGE INSTALLATION

- A training package was developed which can be used in the Copernicus Data Space ecosystem:
 - Create an account, login and start a JNB session
 - Navigate to folder 'mystorage'
 - To install the training package:
 - Create a new folder, e.g. 'TP' and navigate to this folder
 - From the 'Launcher' tab, open a terminal instance and type the following command:

```
wget https://filetransfer.itc.nl/pub/52n/Copernicus_DataSpace/Install_Training_Package.ipynb
```
 - Once the file is downloaded in the folder 'TP', execute the notebook 'Install_Training_Package.ipynb'





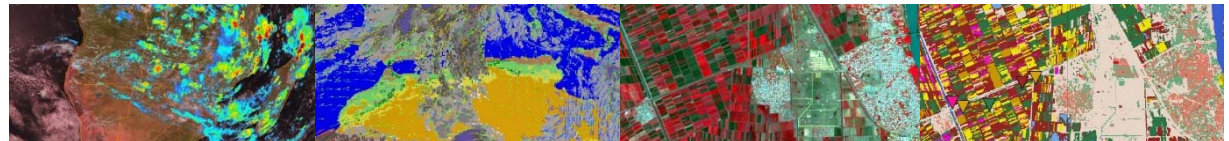
GEO-INFORMATION PROCESSING IN THE CLOUD

INSTALLATION:

- Execute the notebook '**Setup_OpenEO_ILWISPy.ipynb**' to install other required resources

DEMO: navigate to folder >'Demo_SummerSchool'

- Execute the notebook '**openeo_first_steps.ipynb**' to see if you can successfully process Sentinel-2 data
- Execute the notebook '**Wekeo_SST.ipynb**' to see if you can successfully process SST data retrieved from the WEkEO DIAS
- Execute the notebook '**gauge_cor_radar_ilwispy.ipynb**' to see if you can retrieve and process data obtained from external (KNMI) resources





GEO-INFORMATION PROCESSING IN THE CLOUD

Working with sample data from your own region

Setup the resources:

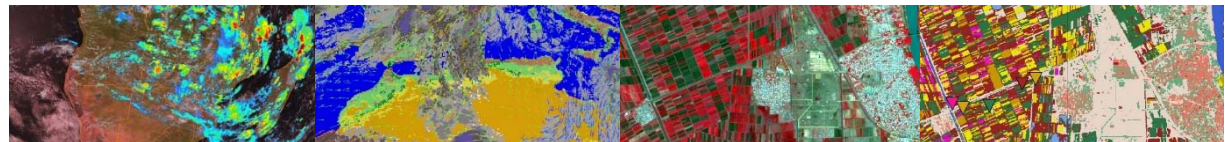
- Execute the notebook [**'Setup_OpenEO_ILWISPy.ipynb'**] to install required resources
- Create a folder and navigate to the folder created
- Open the launcher and type the following command:

```
wget https://filetransfer.itc.nl/pub/52n/Central_Asia_data/asia_sampledata.ipynb
```
- Once the file is downloaded in your folder execute the notebook:
'asia_sampledata.ipynb'

- Meta-Data description: **'data_sources_CA_data.pdf'**



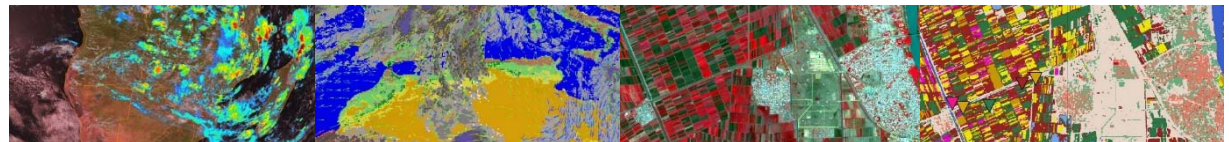
UNIVERSITY OF TWENTE.





GEO-INFORMATION PROCESSING IN THE CLOUD

- Considerations when using the **Copernicus Data Space ecosystem** at a later instance!
 - *All data saved in 'mystorage' will be preserved but all other files will be deleted and additional packages (dependencies) which were installed using pip need to be reinstalled.*
 - *Files in 'mystorage' are kept for up to 15 days from your last login, and you will receive a notification to log into the Jupyter Hub to reset the timer and keep the data preserved. If you do not log in, then after 15 days your files will be deleted.*
 - *Before running notebooks available in the training package, first run the notebook **'Setup_OpenEO_ILWISPy.ipynb'** to install other required resources*





GEO-INFORMATION PROCESSING IN THE CLOUD

Some concluding remarks:

- With resources provided you should be able to adapt and develop your own training materials:
 - *All resources for free – like the open software tools as well as the cloud platform and data;*
 - *Multitude of online data and scripting examples, you can develop training resources according to your curriculum requirements;*
 - *Use of Python and JNB is becoming a standard in academic teaching and training;*
 - *Coding and scripting should be part of the student's curriculum as this is becoming a requirement by the Geo-Information labour market.*

