



SDGs4Namibia

Assessment of SDG Indicators utilizing Satellite Remote Sensing

Insa Otte, Pawel Kluter, Steven Hill, Sebastian Förtsch, Michael Thiel

Department of Remote Sensing, Institute of Geography and Geology, University of Würzburg

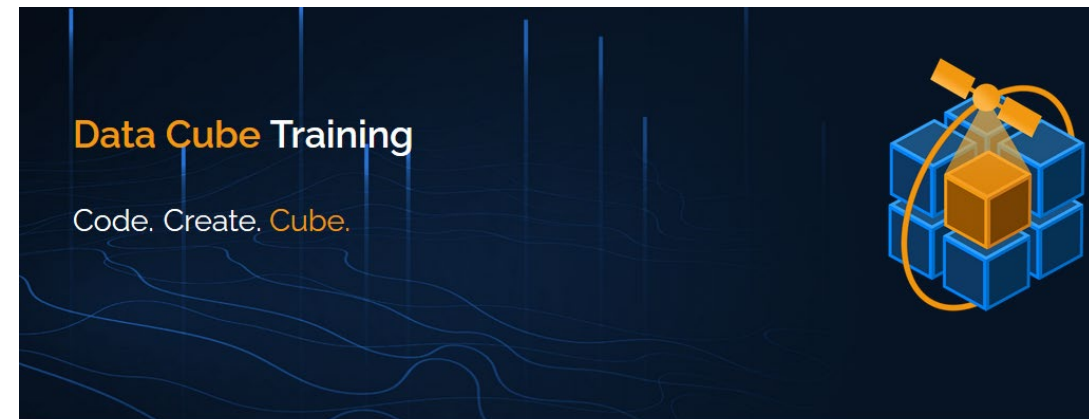
Data Cube - Training

Aim:


- Capacity building for our project partners and beyond

We (will) provide:

- GitHub repository for source code
- Documentation, tutorials and usecases: Jupyter Notebook
- Cube environment with decision ready analysis



Data Cube Training
Code. Create. Cube.




Github

Looking for a useful repository?
Check out our source code.



Docs

In need of tutorials and guides?
Our documentation might help.

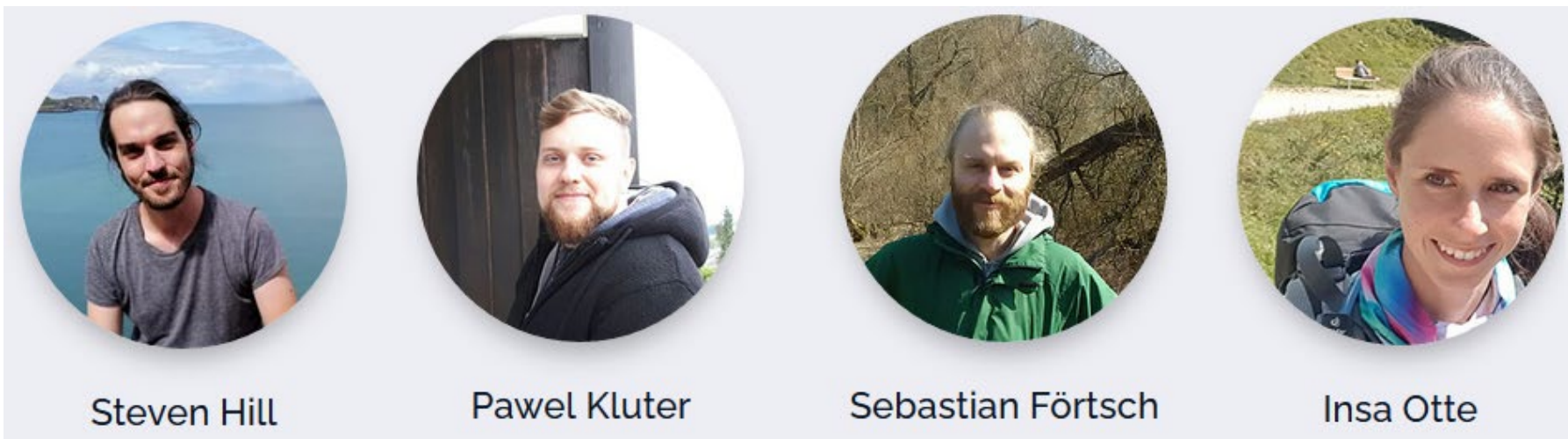


PhenoCube

Ready to see a cube in action?
Visit our live cube environment.



Data Cube Training Team



Steven Hill

Pawel Kluter

Sebastian Förtsch

Insa Otte

	Field of expertise
Steven	tech-coordination, IT infrastrucutre
Pawel	Website, tech/server specialist, ingestion
Sebastian	Jupyter Notebooks, downloads
Insa	Coordination, scientific communication
Kemeng	Introduction into Juypter and Python
Ka Hei	Water change detection Sentinel-2

Data Cube Students

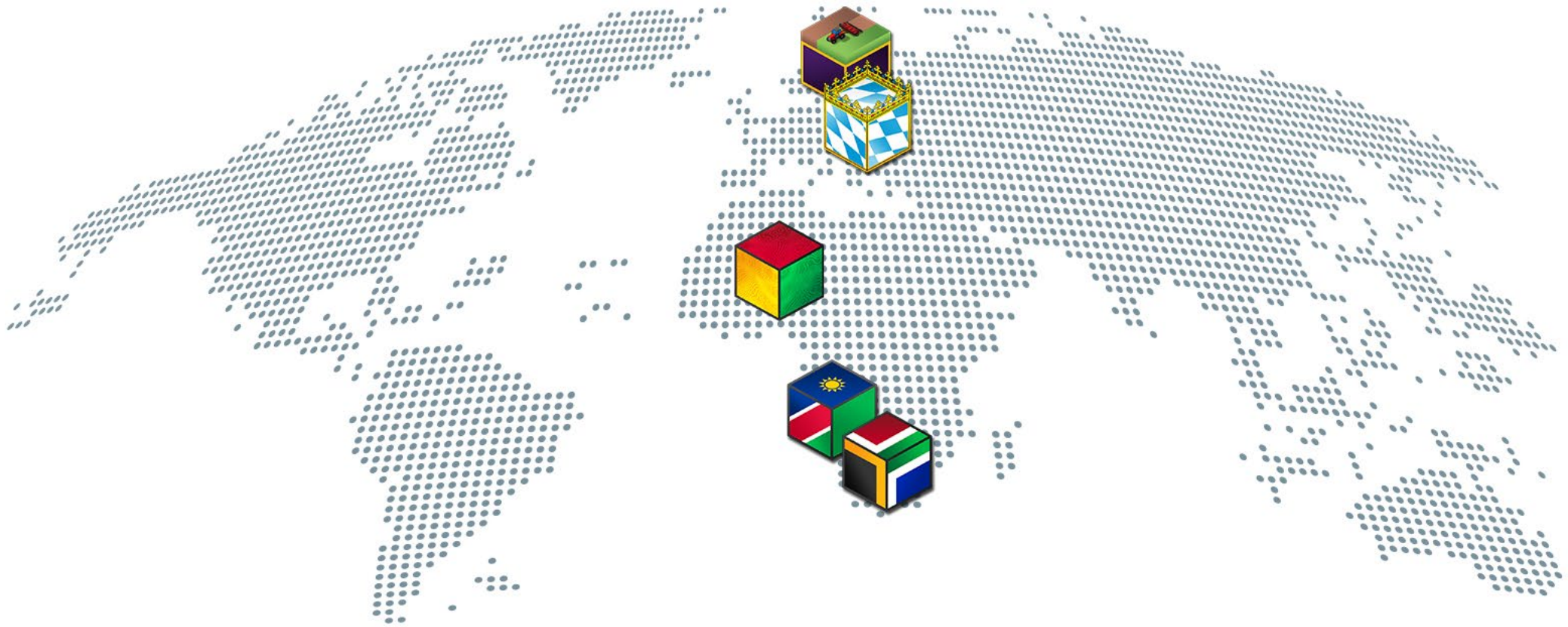


Kemeng
Liu



Ka Hei
Chow





SDGs4Namibia DC Training Schedule:

	Wednesday 12 th May	Friday 14 th May	Monday 17 th May	Tuesday 18 th May	Wednesday 19 th May
10 – 12am	Introduction Data Cube in general [IT]	Hands on: Python & Intro Jupyter Notebooks [GIS]	xarray I: search and load data data structure, indexing, basic plots [GIS]	Spatial analysis [GIS]	SDGs use case 1st run: local 2nd run: DC [GIS]
2 – 4pm	Ingestion: Bring your data in the cube [IT]	Hands on: Intro into EO2Cube [GIS]	xarray II: adv. indexing, statistics + plotting [GIS]	Water detection vegetation [GIS]	Q & A Wrap up [all]



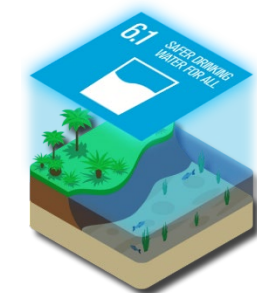
Assessment of SDG Indicators utilizing Satellite Remote Sensing

Pilot project:

Set up of EO Data Cube to evaluate selected SDGs for Namibia
in close cooperation with the NSA (Namibian Statistics Agency)

The main goals are:

- Set-up of an EO data cube (Sentinel-2) for complete Namibia
for effective data handling and analysis (Sentinel-2 data for 2015 and 2020)
- Capacity development measures for data cube handling and EO-data processing
- Implementation of the workflow for one exemplary SDG indicator into the Data Cube:
Indicator 6.6.1: Change in the extent of water-related ecosystems over time

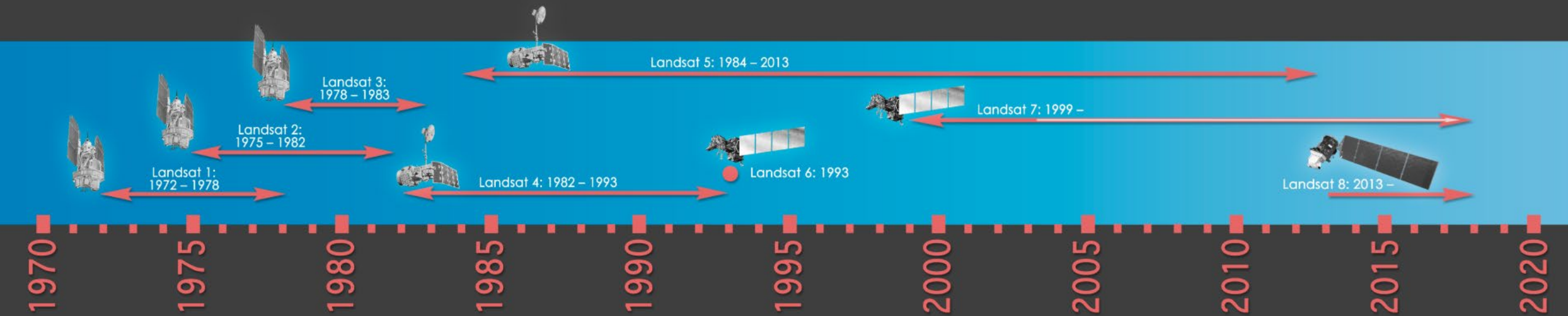


The aim of the SDGs4Namibia data cube is to establish an **earth observation infrastructure for land degradation processes** providing

- uncomplicated EO data access for method development (analysis ready data)
- flexibility and standardization in EO data management
- dynamic working EO tools for research teams and users
- decision-ready EO products
- highly spatio-temporal resolved time-series

Earth observation data

40 Years of Global Observations



(NASA, 2019)

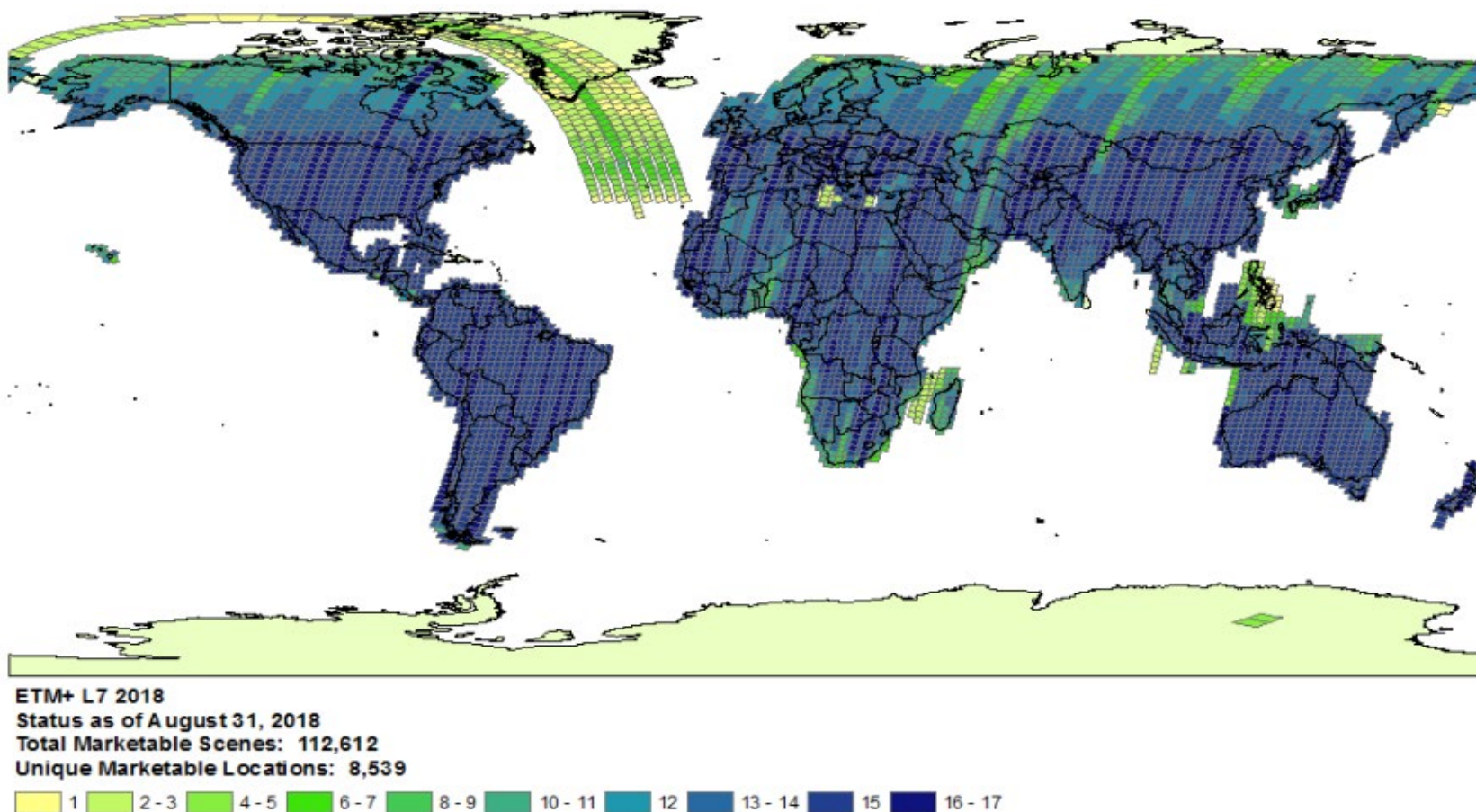


Landsat has already taken more than 7.5 million images of the Earth surface



Credit: United States Geological Survey (Landsat-8 Illustration Above Earth)

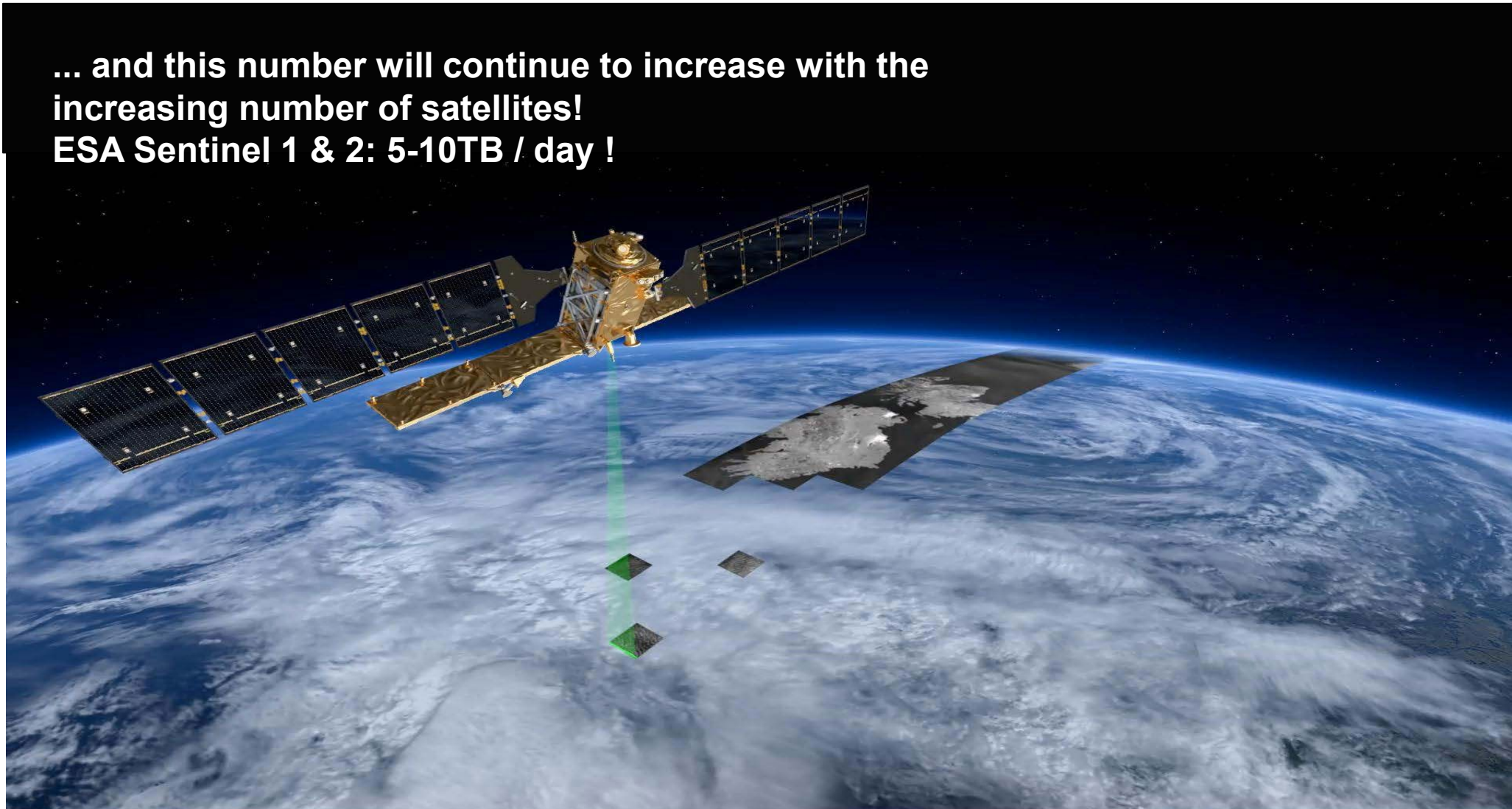
This corresponds to a data volume of **> 7.5 petabytes...**



Source: USGS, 2019



**... and this number will continue to increase with the
increasing number of satellites!
ESA Sentinel 1 & 2: 5-10TB / day !**



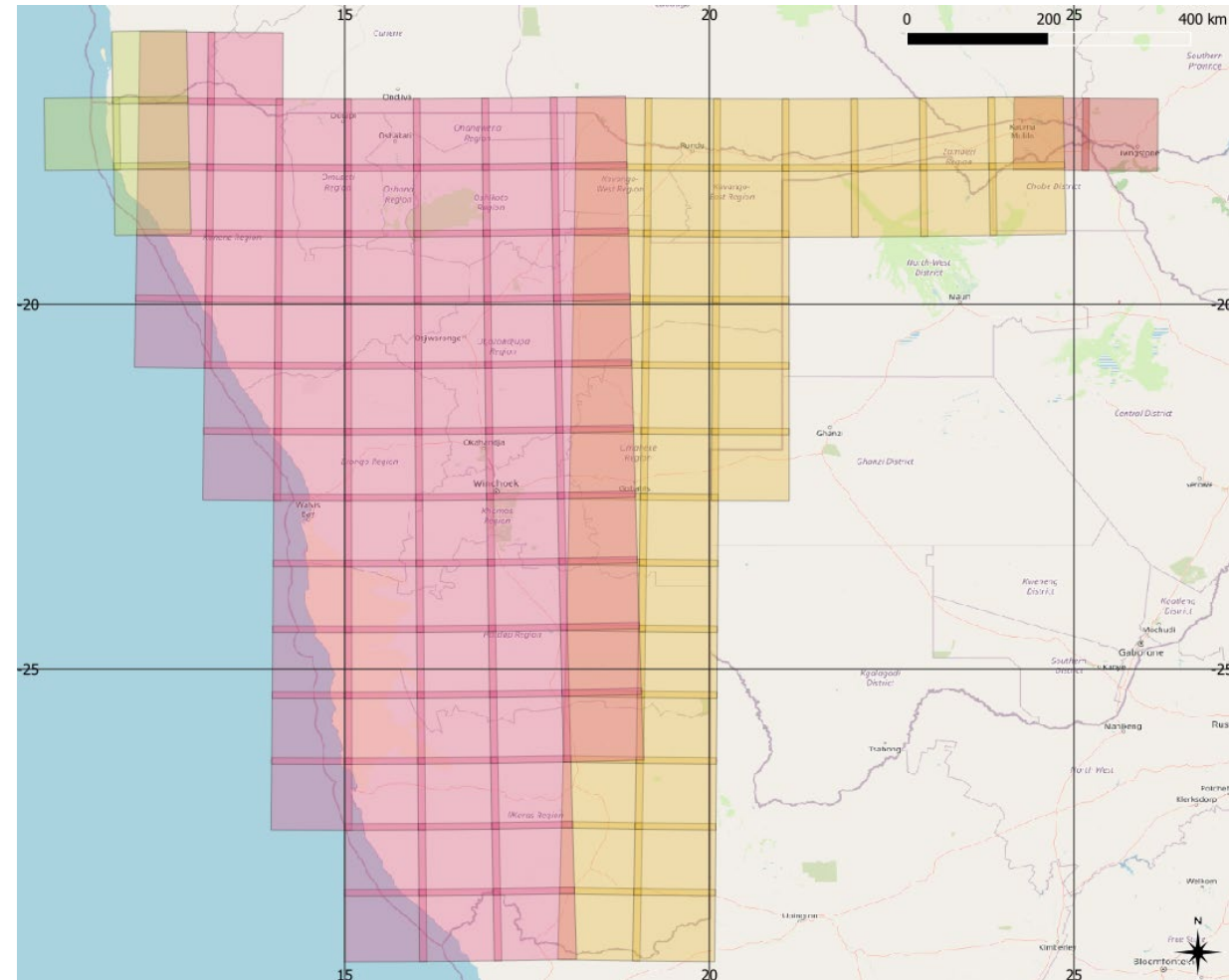
Sentinel-2 > 13,795 scenes for 2020

Full coverage of Namibia

118 tiles for 4 UTM zones

Projection

WGS84 / UTM zone 33S

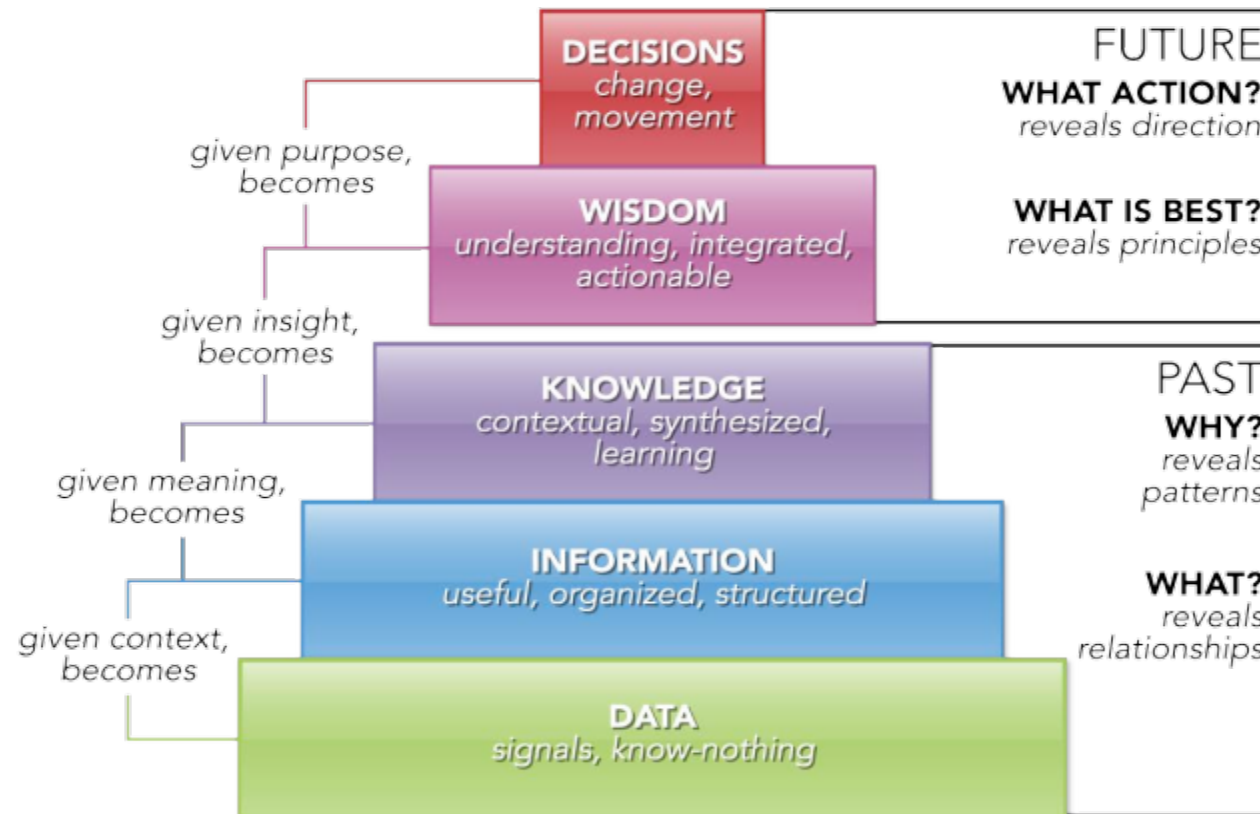


The data deluge



... and how to handle all the data?

How to transform this large amount of data into useful information and support evidence-based decisions



The Future: Open Science

Focus on the results based on open science

- Open Data
- Open Notebooks
- Open Access
- Open Source
- Social Networks
- Citizen Science
- Open educational resources



Sharing knowledge

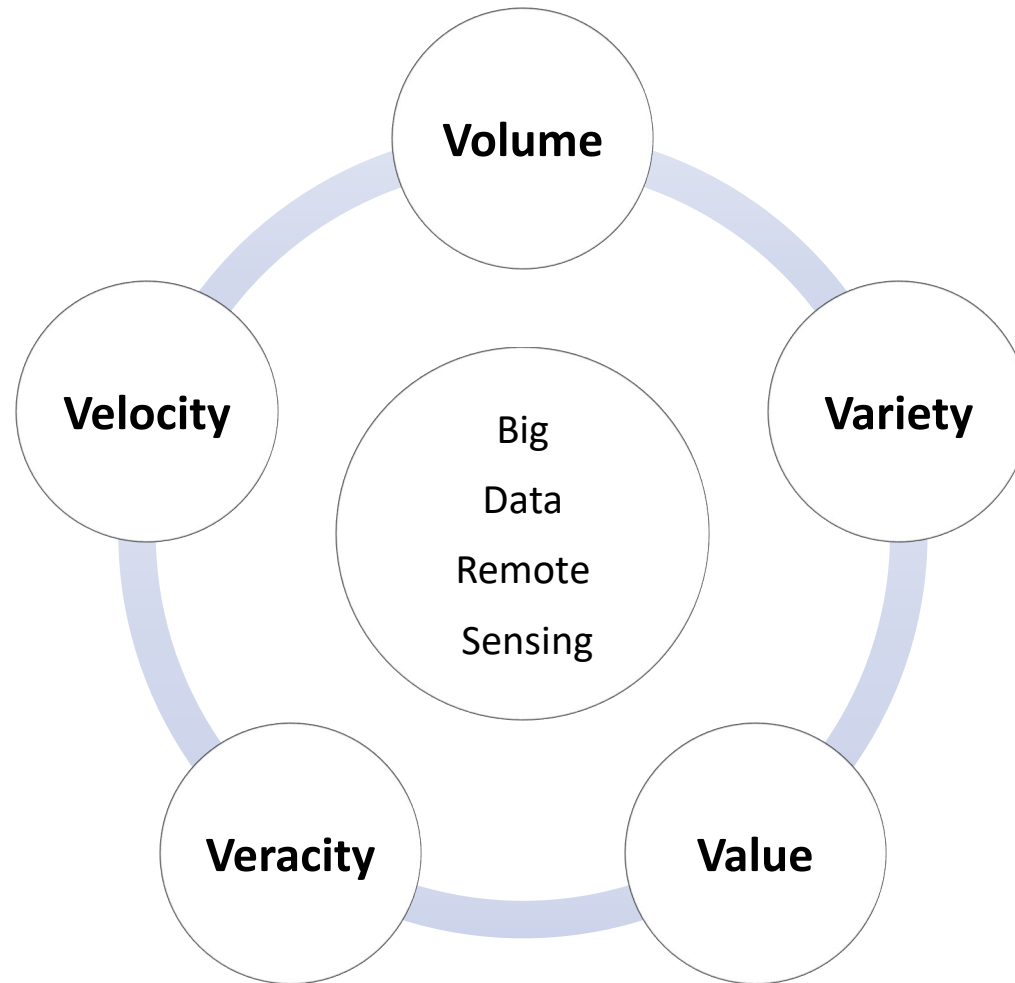
...is becoming easier using the cloud

Mindware, not software!



Courtesy of GEO



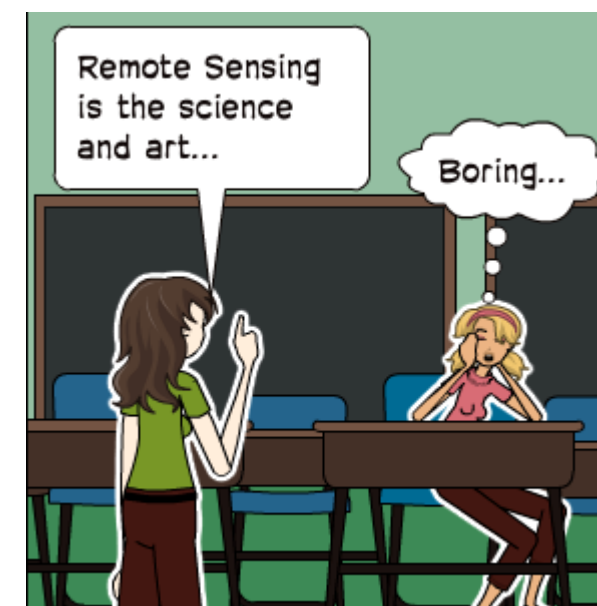


Big Data Challenges:

- Collect
- Manage
- Store
- Archive
- Analysis
- Visualize
- Distribute

Why do so few people use satellite data?

- Scientific knowledge is required to understand... (what kind of data? / resolution? / what type?)
- Difficult to access and download
- Difficult to prepare... atmospheric corrections, alignment, formats
- Challenging analysis
- Need for training and capacity building



<https://www.kejoyce.com/blog/>

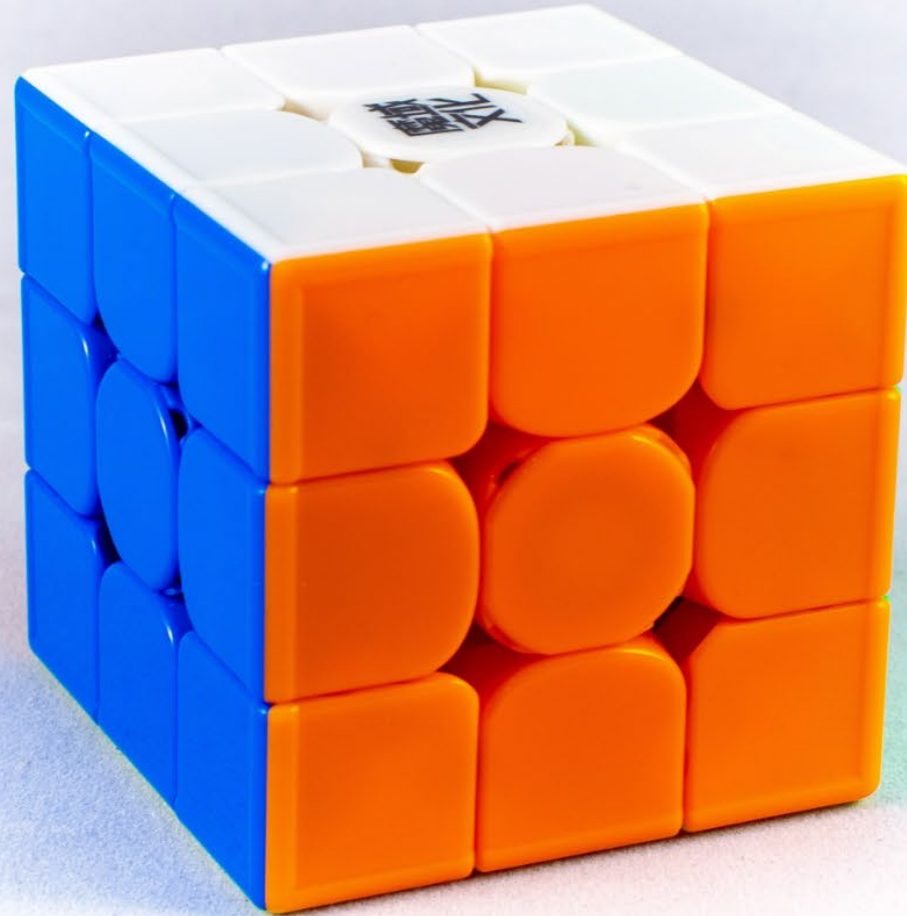


Innovative data analysis infrastructure for the analysis of earth observation
data in order to support decision making

A new solution... **DATA CUBES?**

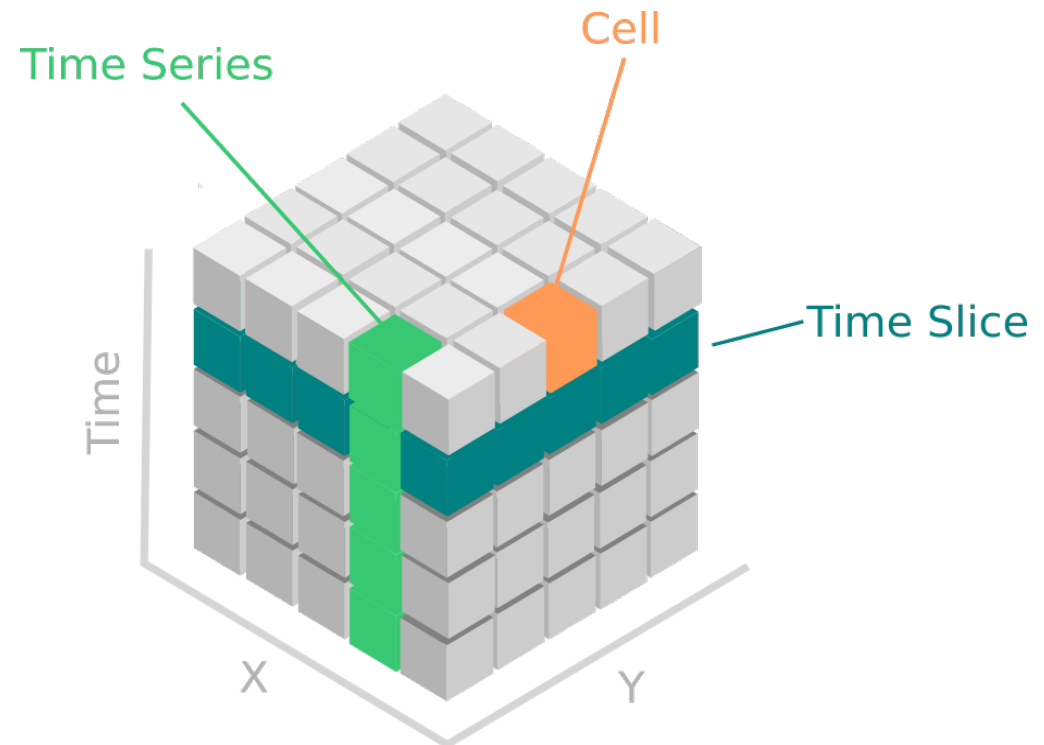


What is a data cube ?



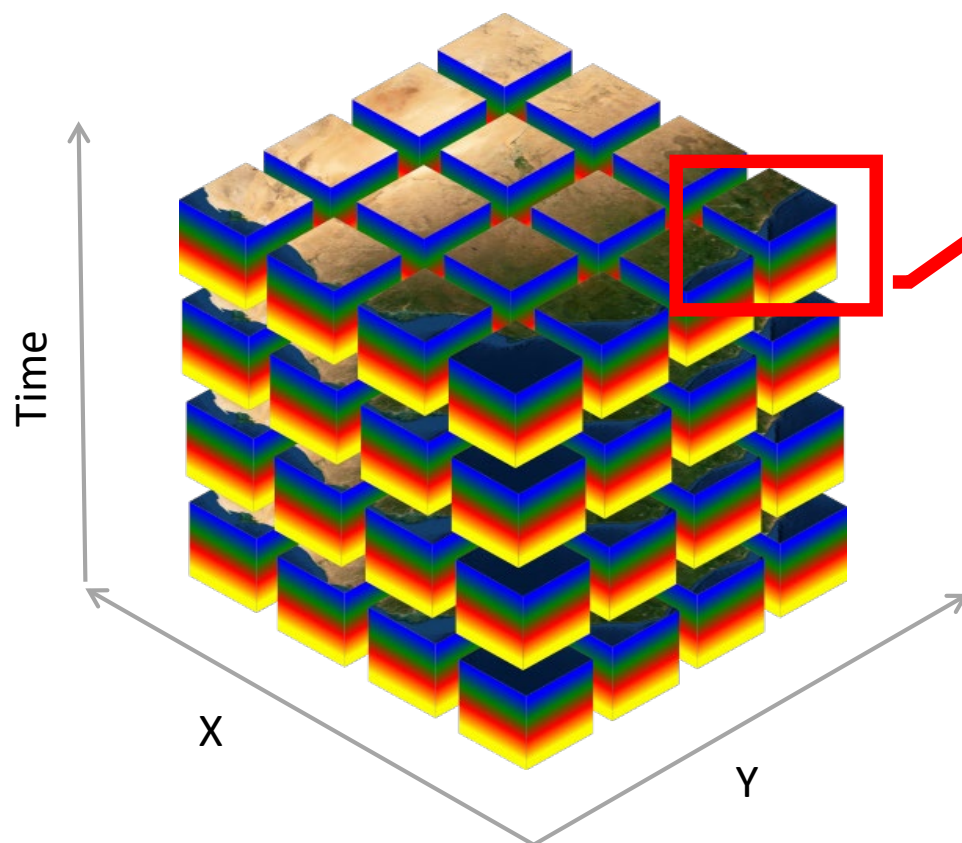
What is a data cube ?

- Used to represent data along some measure of interest
- Can be 2-dimensional, 3-dimensional, or higher-dimensional
- Each dimension represents some attribute in the database
- Each cell in the data cube represents the measure of interest



Earth observation data cubes

Multi-dimensional stack (space, time, data type) of spatially aligned pixels & used for efficient access and analysis



Sentinel-2 (NetCDF)

Spatial: 10 km x 10 km

Time: 2020-01-22 22:10:05

Bands: 12

What is Open Data Cube?

The Open Data Cube Infrastructure provides an integrated gridded **data analysis environment** providing **long-term analysis** ready earth observation data from **multiple satellites and other acquisition systems**.

Why using this approach?

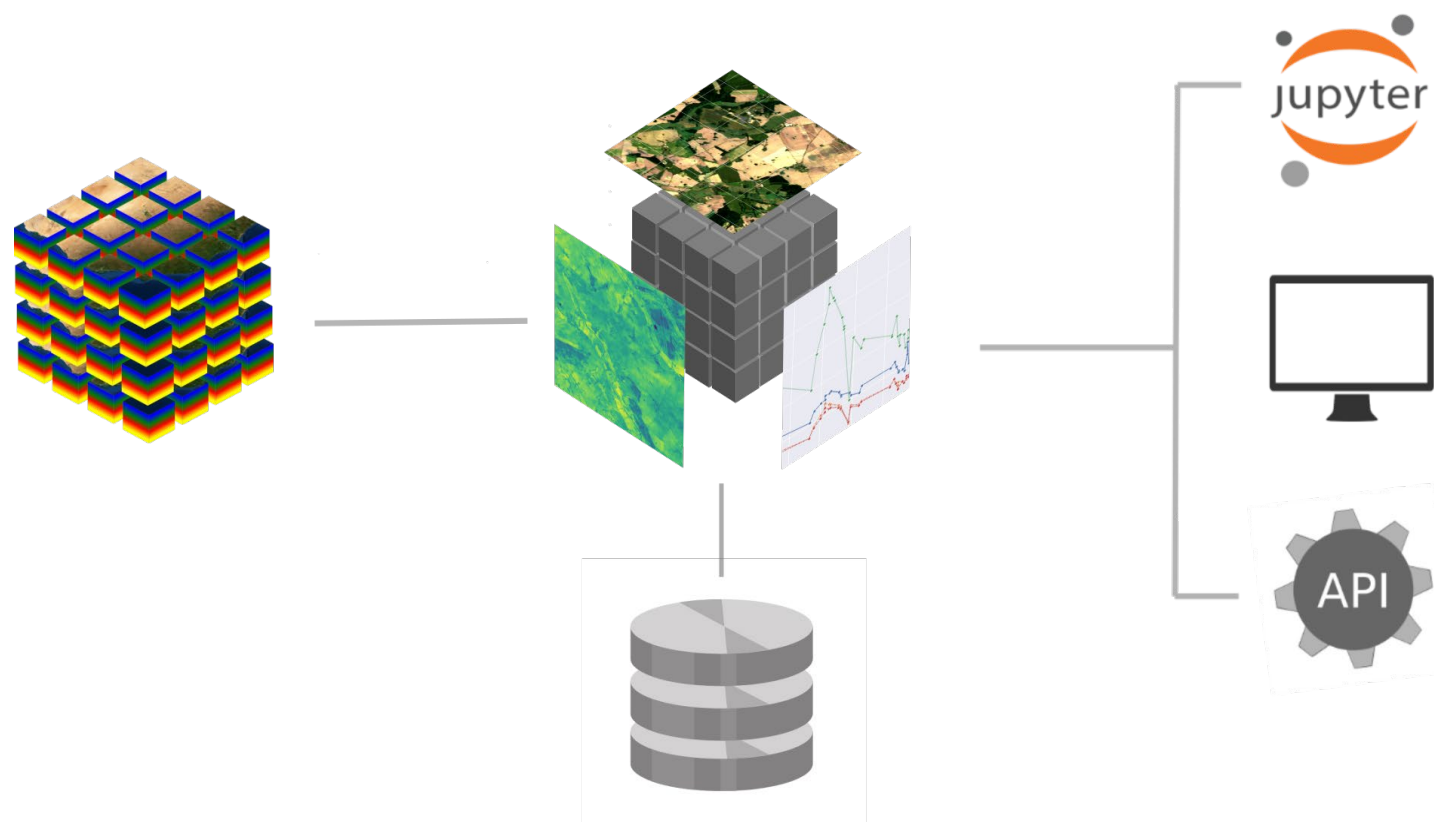
The Data Cube framework allows multi-sensor environmental monitoring on a **cloud-based platform** for an **efficient and user-oriented analysis** of land surface phenomena based on **multi-temporal earth observation data**.



OPEN DATA CUBE

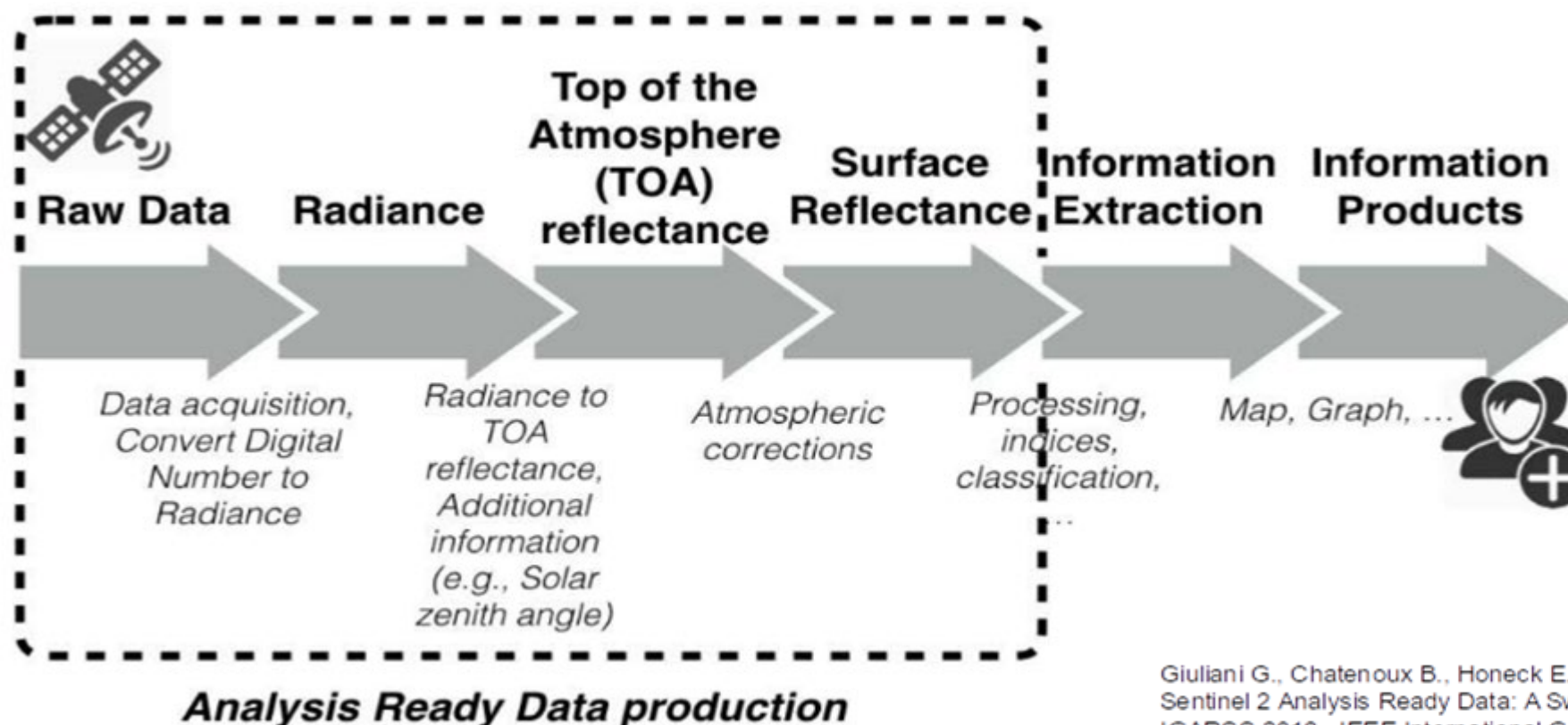
The idea of the SDGs4Namibia data cube is ...

to **increase the value and impact of global Earth observation satellite data for Namibia** by providing an accessible exploitation architecture for an **efficient and user-oriented analysis based** on multi-temporal earth observation data.



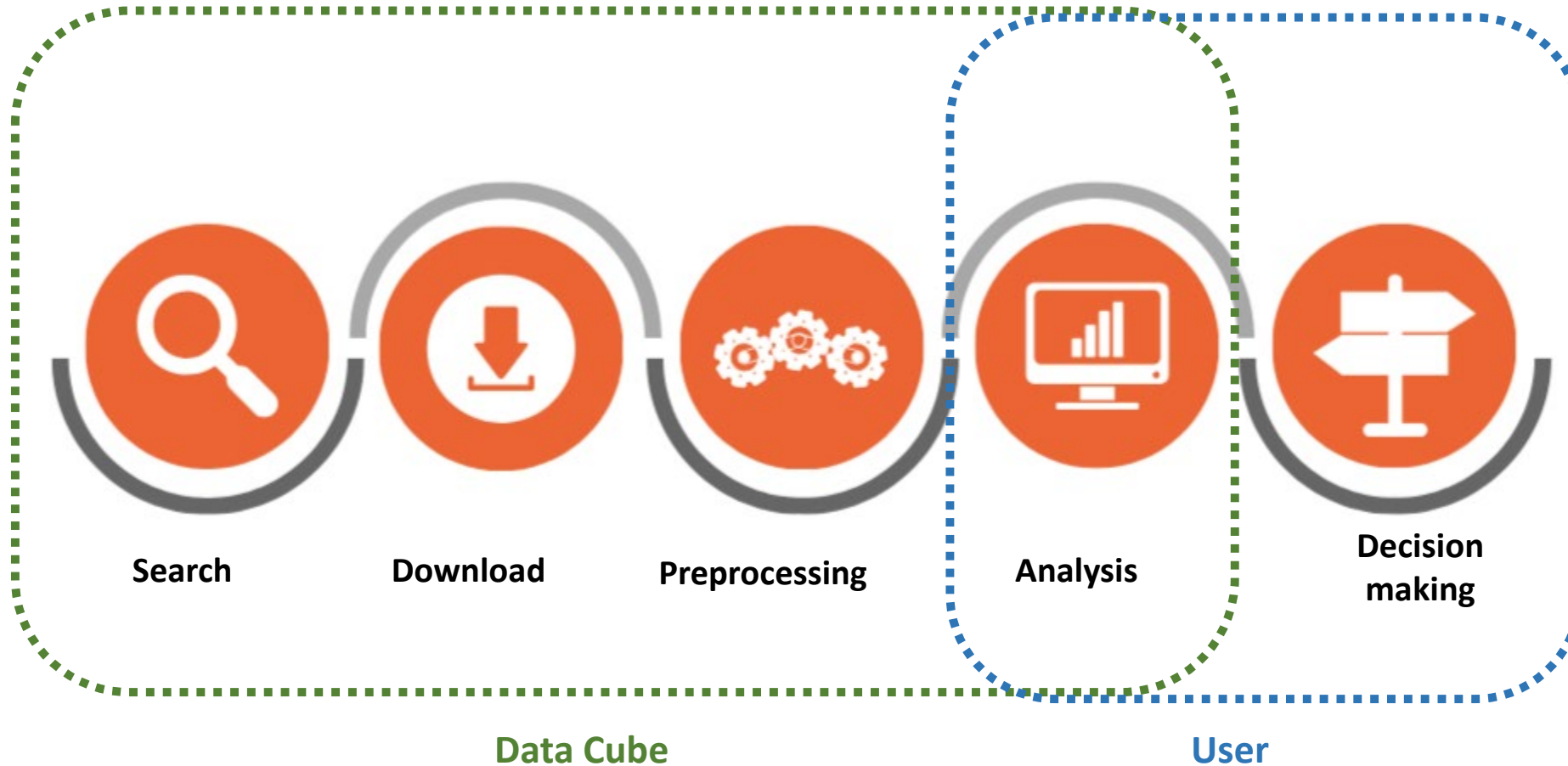
Analysis Ready Data are key to reduce the burden on EO data users

Spending more time in analyzing data than searching & pre-processing data...

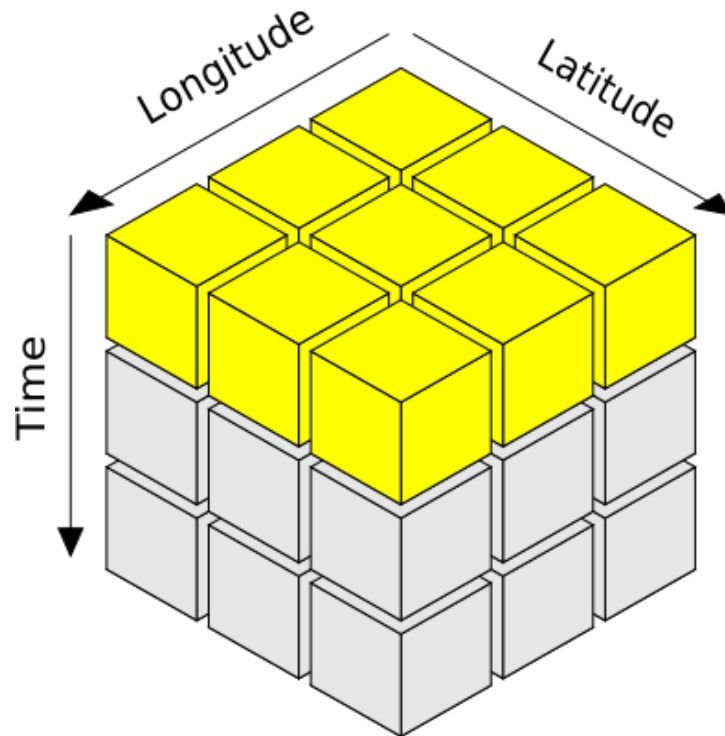


Giuliani G., Chatenoux B., Honeck E., Richard J.-P. (2018) Towards Sentinel 2 Analysis Ready Data: A Swiss Data Cube Perspective. In: IGARSS 2018 - IEEE International Geoscience and Remote Sensing Symposium. Valencia (Spain). p. 8668-8671

Workflow data cube



Sampling a data cube



A single time slice,
similar to a standard „scene“
can be used to assess a single point in time

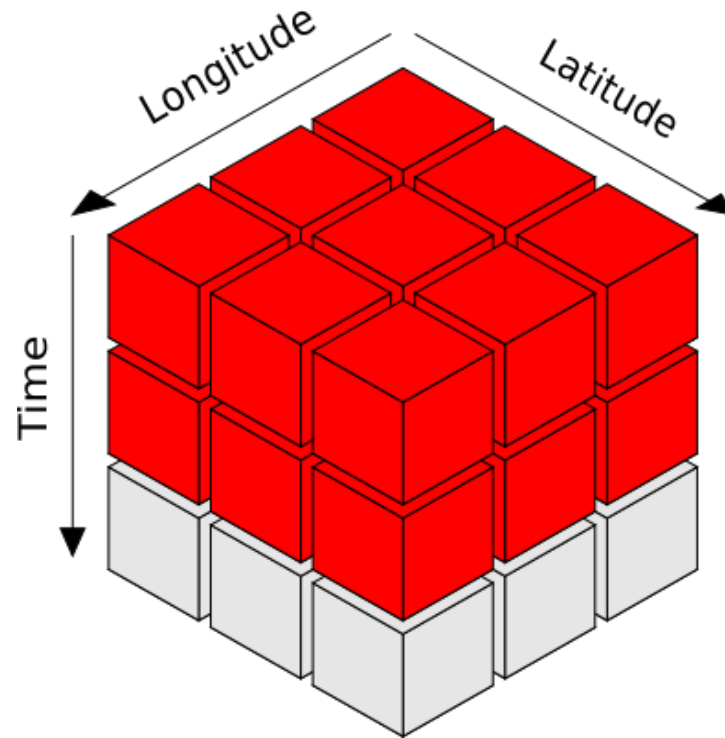
Pixels in a data cube are processed, aligned
and compressed and ready for data analysis

All available data at one point as time

Large scale mosaic



Sampling a data cube



Several time slices can be combined into one to form a „**mosaic**“. This is often used to reduce clouds or create seasonal or annual images.

Typical **mosaics**:

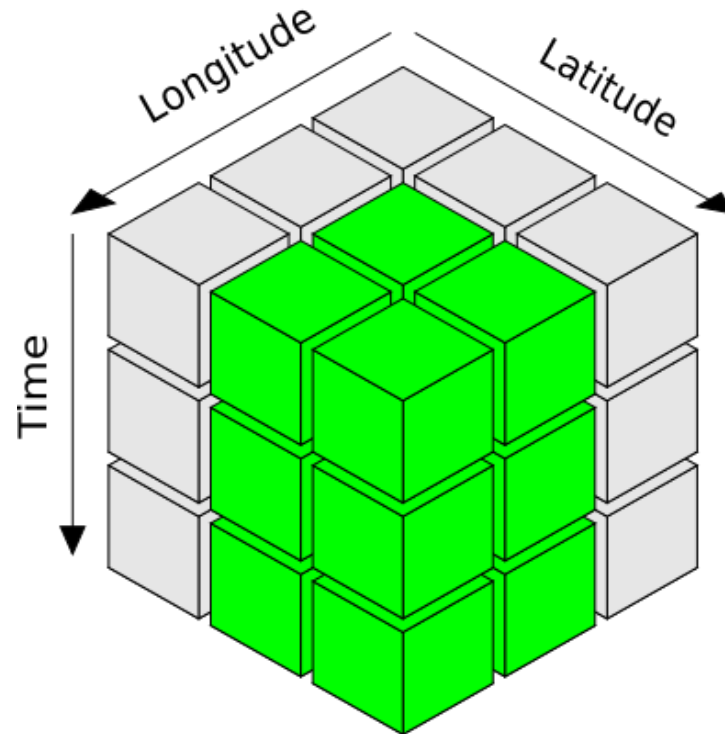
- Most/least recent pixel
- Mean/median
- Geomedian
- Min/max NDVI

Combine several time slices

Cloud removal



Sampling a data cube



Examples of time series analyses include:

- Land change (PyCCD)
- Water change (WOFS)
- Parameter variation along a transect (Hovmoller plot)

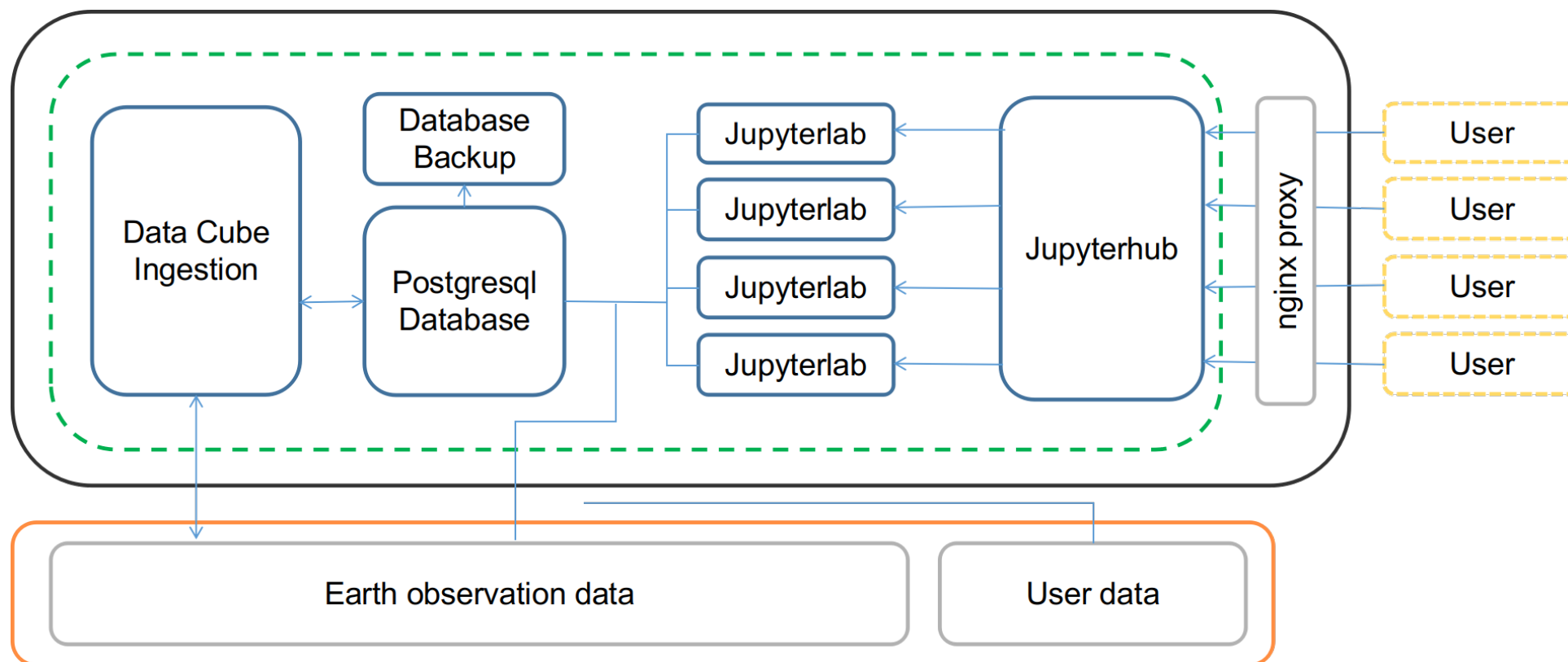
Time series analyses consider the variation of data over time to assess change

All slices for a specific region

Time series analysis



SDGs4Namibia data cube environment



Virtual Machine



NFS Storage



Docker network





Docker container

How do we access our SDGs4 Namibia data cube?



01_jupyter_introduction.ipynl X

+
-
Copy
Paste
Run
Stop
Refresh
Markdown
Python 3

Introduction to Jupyter notebooks

- [Sign up to the JupyterHub](#) to run this notebook interactively from your browser
- **Compatibility:** Notebook currently compatible with the Open Data Cube environments of the University of Wuerzburg
- **Prerequisites:** There is no prerequisite learning required.

Background

Access to implementations of the [Open Data Cube](#) such as [Digital Earth Australia](#) and [Bavarian Data Cube](#) is achieved through the use of Python code and [Jupyter Notebooks](#). The Jupyter Notebook is an interactive web application that allows viewing, creation and documentation of live code. Notebook applications include data transformation, visualisation, modelling and machine learning.

What is Jupyter?

The *Jupyter Project* is an open source effort that evolved from the IPython project to support interactive data science and computing. Besides `Python`, it also supports many different programming languages including `R` and `Julia`.

(If you're familiar with the `R` programming language, Jupyter Notebook can be compared to R Markdown).

Jupyter is an open source platform that contains a suite of tools including:

- **Jupyter Notebook: A browser-based interactive development environment (IDE) that allows users to write and run e.g. `python` codes in individual cells where the output is displayed under each executed cell.**



How do we access our SDGs4 Namibia data cube?



01_jupyter_introduction.ipynl X

Python 3

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

eo2cube

Introduction to Jupyter notebooks

- [Sign up to the JupyterHub](#) to run this notebook interactively from your browser
- **Compatibility:** Notebook currently compatible with the Open Data Cube environments of the University of Würzburg
- **Prerequisites:** There is no prerequisite learning required.

Background

Access to implementations of the Open Data Cube such as [Bosnian Data Cube](#) and [Bavarian Data Cube](#) is achieved through the use of Python code and [Jupyter Notebooks](#). The Jupyter Notebook is an interactive web application that allows viewing, creating and editing of live code. Notebook applications include data transformation, visualisation, modelling and machine learning.

What is Jupyter?

The [Jupyter Project](#) is an open source effort that evolved from the IPython project to support interactive data science and computing. Besides [Python](#), it also supports many different programming languages including [R](#) and [Julia](#).

(If you're familiar with the [R](#) programming language, Jupyter Notebook can be compared to R Markdown).

Jupyter is an open source platform that contains a suite of tools including:

- **Jupyter Notebook:** A browser-based interactive development environment (IDE) that allows users to write and run e.g. [python](#) codes in individual cells where the output is displayed under each executed cell.

Friday 10 – 12 am:
Python & Intro Jupyter Notebooks



What will be inside the SDGs4Namibia data cube?

Remote sensing dataset

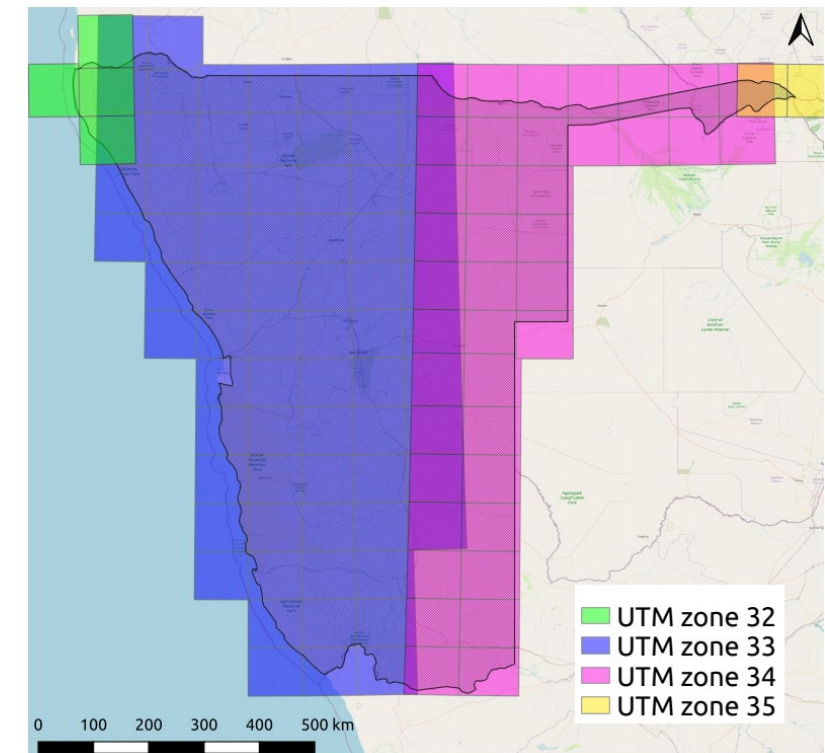
- ✓ Level 2A Sentinel-2A and Sentinel-2B (10 m, 20 m, 60 m)
→ 2015 & 2020 is ingested
- ✓ DLR's GWP for 2013, 2018 (MODIS based)

Preprocessing

- ✓ Atmospheric correction: sen2cor
- ✓ Cloud mask: sen2cor

Products

- ✓ GWP - derive SDG indicator 6.6.1



Take away & Outlook

- Analysis ready data sets are key to reduce the burden on EO data users
- Data cubes can provide the long baseline required to determine trends, define present and inform future
- Visit our website eo2cube (<http://datacube.remote-sensing.org/>)



Thanks for listening!!

SDGs4Namibia DC Training Schedule:

	Wednesday 12 th May	Friday 14 th May	Monday 17 th May	Tuesday 18 th May	Wednesday 19 th May
10 – 12am	Introduction Data Cube in general [IT]	Hands on: Python & Intro Jupyter Notebooks [GIS]	xarray I: search and load data data structure, indexing, basic plots [GIS]	Spatial analysis [GIS]	SDGs use case 1st run: local 2nd run: DC [GIS]
2 – 4pm	Ingestion: Bring your data in the cube [IT]	Hands on: Intro into EO2Cube [GIS]	xarray II: adv. indexing, statistics + plotting [GIS]	Water detection vegetation [GIS]	Q & A Wrap up [all]

