Getting started with Satellite Data in Python

EOX IT Services

PyLadies Vienna

8. 3. 2024, Women&&Code

Mentorship Programme

Meetups

Python Beginners courses

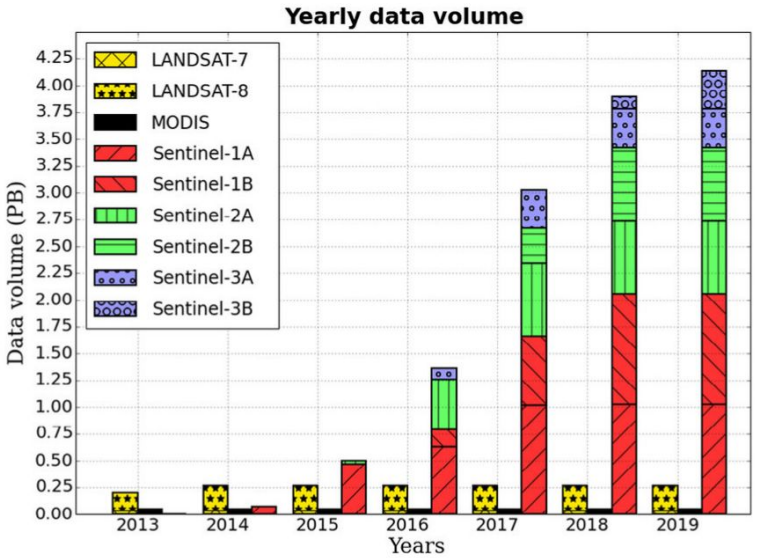
Application close tonight!

PyLadies Vienna
pyladies.at

What are EO data and why should we care

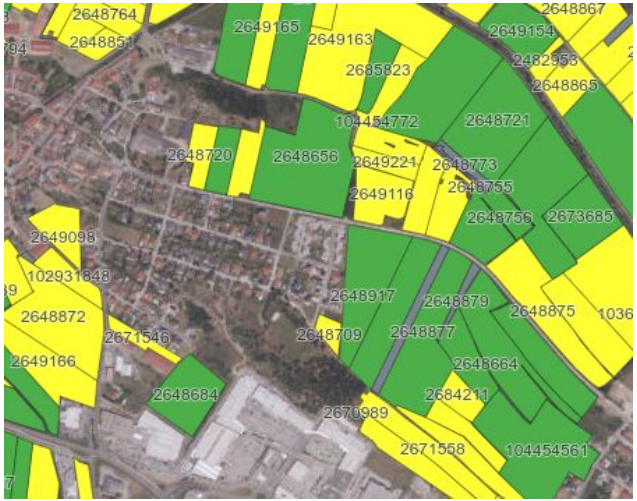
- Data captured from Space
- New missions every year
- Only Optical data here today - Sentinel 2
- What are bands?
- Many different types:
    - different resolution
    - different processing levels
    - different instruments
    - different revisit times
    - different bands available

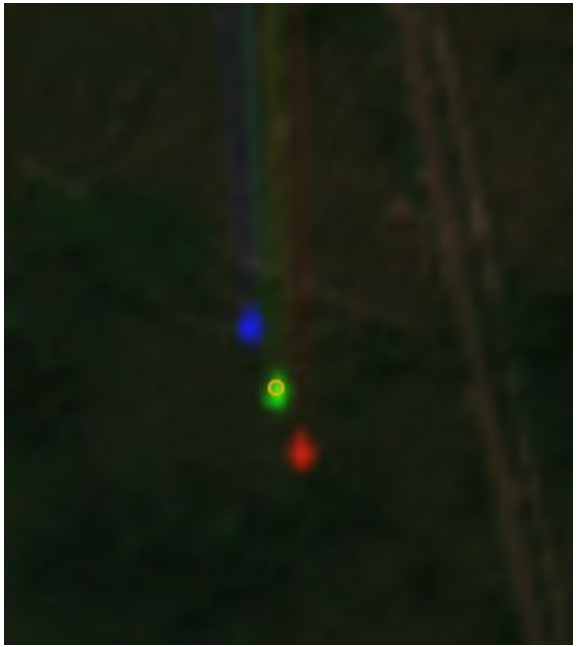
- But everything is basically numpy array



Possible applications

- Land classification
- Crop type monitoring
- Sea ice monitoring
- Water events and changes
- Soil monitoring
- Forest monitoring
- Plane detection
- …





Where and how to find your Data

- Web catalog clients with a map interface and filtering
- Direct API access - search queries
  - STAC, OpenSearch, OData, CSW

Metadata and Properties

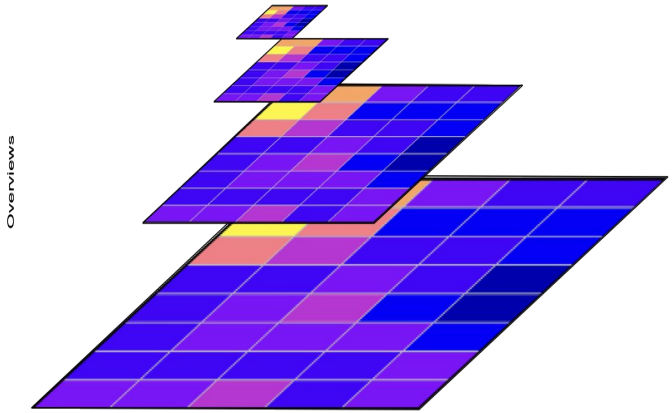
- Filter using metadata
- Bounding box (spatial)
- Time interval
- Maximum cloud coverage

Platforms

- Copernicus Data Space Ecosystem
  - official access point by ESA
- the Big Five host public Sentinel 2 data as well

Data formats - size vs utility

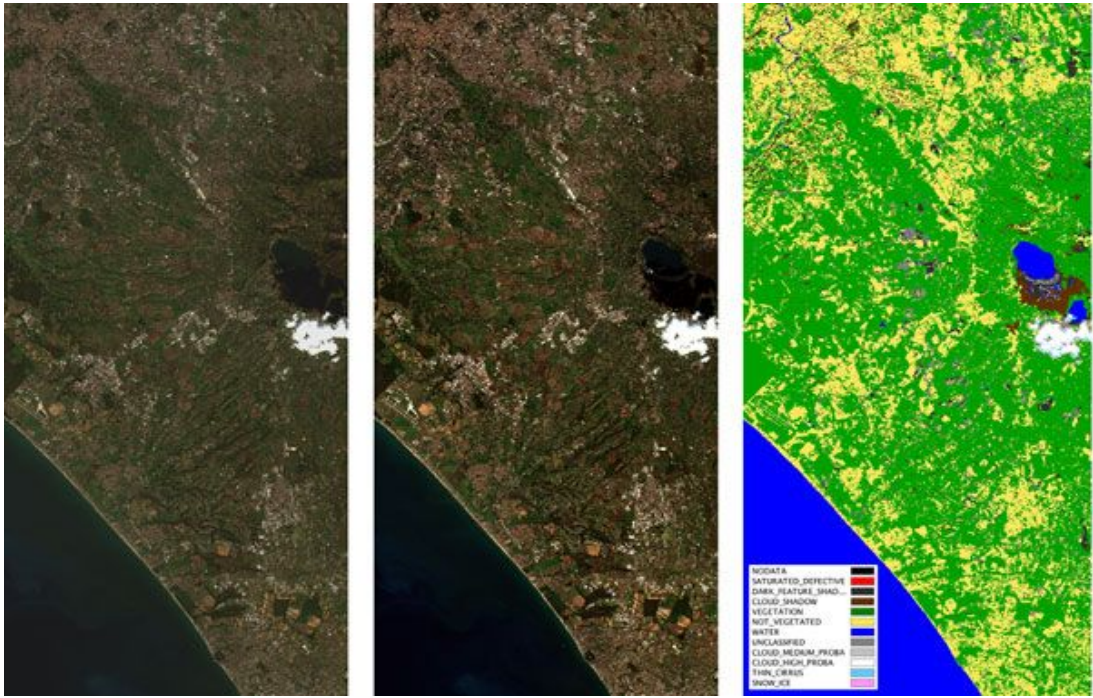
- Cloud optimized formats for read access
  - overviews, header allowing range HTTP requests
- GeoTIFF, COG
- SAFE (Standard Archive Format for Europe)
- JPEG2000
- HDF5



Data Processing

- Raw data from satellite are usually pre-processed to different levels
- Different levels of corrections
  - Atmospheric correction
  - Radiometric correction
- GDAL as underlying tool
- → Python bindings
- Rasterio

- Reprojection
- Clip by geometry
- Merge bands
- Format exchange
- Mosaic creation
  → AR data often available



EO Data Analysis

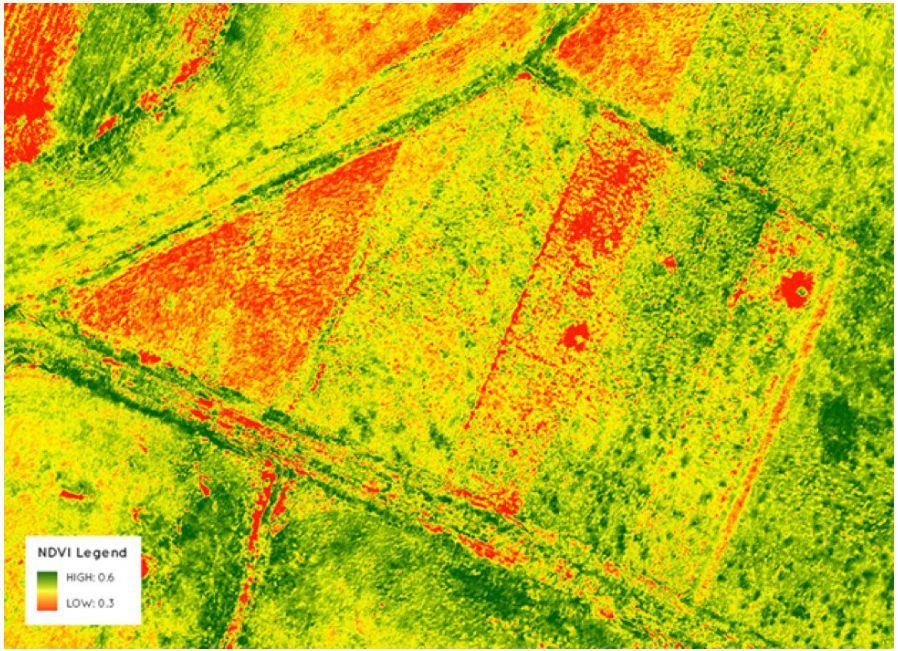
- ML models based on standard architectures
- Mainly classification and detection task
- “big data” processes - large volumes and large scales even using small areas
- Problem of never enough training data and labelling
- Nodata values - common pitfall

But:

- For many task indexes are used
- …and they are (still) in many cases enough!
- Enhancement with another data - census data, energy grid, …

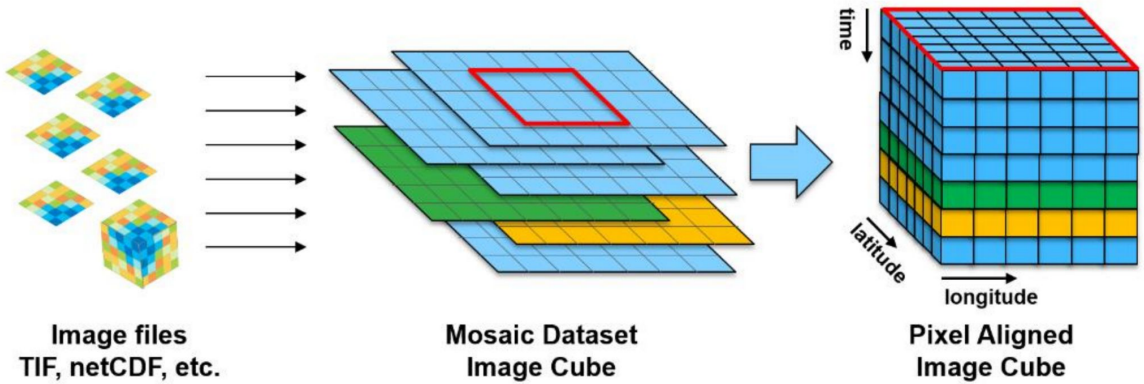
NDVI Index

- Normalized Difference Vegetation Index
- (NIR - RED) / (NIR + RED)
- In theory from -1 to +1
- In practice - close to zero means no vegetation
- Harvest detection



Datacubes

- Not only data storage in EO world
- Data are together with temporal dimension
- Different data information put together



Visualization

- UInt16 -> RGB conversion for display
- Different band combinations to highlight phenomena
- Map overlays or background maps (mosaics)

Summary

- Satellite data are cool!
- Play EOGuesser and share your score on social media :) https://s2maps.eu/eo-guesser/
- If you want to know more, just find me (I can talk about EO data for hours)
- New free MOOC just launch - Cubes and Clouds
- We are looking for a DevOps engineer to join me in a team